

UNIVERSIDADE FEDERAL DO PARANÁ

ESTIMATIVA DA ALTURA DA CAMADA LIMITE
ATMOSFÉRICA COM USO DE REDES NEURAIIS
ARTIFICIAIS

Isabela Godoy Cabral

Curitiba
2015

UNIVERSIDADE FEDERAL DO PARANÁ

ESTIMATIVA DA ALTURA DA CAMADA LIMITE
ATMOSFÉRICA COM USO DE REDES NEURAIS
ARTIFICIAIS

Isabela Godoy Cabral

Trabalho apresentado como requisito parcial à obtenção
do grau de Mestre em Engenharia Ambiental, no
Programa de Pós-Graduação em Engenharia Ambiental,
Setor de Tecnologia, da Universidade Federal do Paraná.

Orientador: Prof. Dr. Ricardo Carvalho de Almeida

Curitiba
2015

C117e

Cabral, Isabela Godoy

Estimativa da altura da camada limite atmosférica com uso de redes neurais artificiais/ Isabela Godoy Cabral. – Curitiba, 2015.

72 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-graduação em Engenharia Ambiental, 2015.

Orientador: Ricardo Carvalho de Almeida .

Bibliografia: p. 70-72.

1. Camada limite planetária. 2. Redes neurais (Computação). 3. Radiossondas. I. Universidade Federal do Paraná. II. Almeida, Ricardo Carvalho de. III. Título.

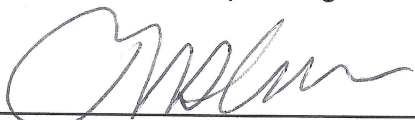
CDD: 551.5130285

TERMO DE APROVAÇÃO

ISABELA GODOY CABRAL

ESTIMATIVA DA ALTURA DA CAMADA LIMITE ATMOSFÉRICA COM USO DE REDES NEURAIS ARTIFICIAIS

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Engenharia Ambiental, Setor de Tecnologia, da Universidade Federal do Paraná, pela seguinte banca examinadora:



Prof. Dr. Ricardo Carvalho de Almeida
Orientador – Departamento de Engenharia Ambiental, UFPR



Dr. José Eduardo Gonçalves
Instituto Tecnológico SIMEPAR



Prof. Dr. Nelson Luís da Costa Dias
Departamento de Engenharia Ambiental, UFPR

Curitiba, 25 de novembro de 2015

RESUMO

A altura da camada limite atmosférica (h) é um parâmetro fundamental em modelagem atmosférica, sendo a dimensão vertical do domínio em modelos de dispersão de poluentes na atmosfera. As relações entre parâmetros meteorológicos e características da superfície é apenas parcialmente conhecida, implicando em dificuldades na obtenção dessa variável. Nesse trabalho empregamos uma rede neural artificial do tipo *multilayer perceptron*, com algoritmo de treinamento *backpropagation*, para a estimativa de h a partir de dados de superfície, obtidos em estações meteorológicas. A base de dados para treinamento foi desenvolvida para esse trabalho, a partir de perfis de propriedades obtidos por radiossondagens e disponibilizados pelo *Stratosphere-troposphere Processes and their Role in Climate* (SPARC). Foram testadas diversas combinações de dados de entrada, incluindo também parâmetros meteorológicos de altitude. A melhor configuração empregando apenas dados de superfície obteve bons resultados, com $IOA = 0.9432$ e $r = 0.9017$, sendo indicada sua aplicação para estimativa de h em localidades onde não estejam disponíveis informações de perfis de propriedades atmosféricas.

Palavras-chave: Camada Limite Atmosférica, Redes Neurais Artificiais, radiossondagem.

ABSTRACT

The Boundary layer height (h) is considered a fundamental parameter for atmospheric modeling and it is the vertical extension of models domain for pollutants dispersion. Despite its importance, the relations for obtaining h are still only partially understood, limiting the quality of models results. In this work we present the use of a multilayer perceptron artificial neural network for estimating h , applying surface parameters obtained in ground based meteorological stations. The network was trained using a backpropagation algorithm, and it was tested for many different combinations of input data. For that, we needed to develop a hole data base from radiosounding profiles, which were available in the SPARC's (Stratosphere-troposphere Processes and their Role in Climate) website. The best combination showed $IOA = 0.9432$ and $r = 0.9017$. Therefore, operation of an artificial neural network is recommended for estimating h when meteorological parameters profiles are not available.

Keywords: Atmospheric boundary layer, Artificial neural networks, radiosounding.

Lista de Figuras

2.1	Imagem de uma radiossonda com receptor GPS acoplado (a) e de um balão atmosférico desinflado (b). Fonte: NOAA	11
2.2	Representação de um neurônio biológico.	12
2.3	Representação de um neurônio artificial.	13
2.4	Representação de uma rede neural artificial do tipo <i>perceptron</i>	14
2.5	Representação de uma rede neural artificial do tipo <i>perceptron</i> multicamadas.	15
2.6	Representação de uma superfície de erros, contendo um mínimo local e um mínimo global.	17
2.7	Curvas de treinamento e teste, ao longo das épocas de treinamento.	17
3.1	Imagem de satélite da localização da estação de Amarillo. Fonte: API do Google Maps.	29
3.2	Médias ao longo de 30 anos de precipitação e temperatura para Amarillo.	29
3.3	Imagem de satélite da localização da estação de Peachtree. Fonte: API do Google Maps.	30
3.4	Médias ao longo de 30 anos de precipitação e temperatura para Peachtree.	30
3.5	Médias ao longo de 30 anos de precipitação e temperatura para Reno.	31
3.6	Imagem de satélite da localização da estação de Reno. Fonte: API do Google Maps.	31
3.7	Distribuição dos valores de h	36
3.8	Distribuição dos valores de $\ln(h)$	36
4.1	Radiossondagens da estação de Peachtree, às 18 : 30 h no horário local. (a) Radiossondagem do dia 10 de junho 00 UTC. (b) Radiossondagem do dia 15 de junho 00 UTC.	44
4.2	Radiossondagens da estação de Reno, com lançamento às 15 : 30 h no horário local. (a) Lançamento de 1º de janeiro, 00 UTC. (b) Lançamento de 7 de janeiro, 00 UTC.	46

4.3	Radiossondagens com lançamento à tarde, para as estações de Amarillo e Peachtree. (a) Lançamento na estação de Peachtree, dia 23 de dezembro às 18 : 30 h no horário local, 00 UTC. (b) Lançamento na estação de Amarillo, dia 5 de dezembro às 17 : 30 h no horário local, 00 UTC.	47
4.4	Radiossondagens com lançamento à tarde, meses de junho e julho. (a) Radiossondagem da estação de Reno, com lançamento às 15 : 30 h do dia 1º de julho no horário local, 00 UTC. (b) Radiossondagem da estação de Reno, com lançamento às 15 : 30 h do dia 10 de junho no horário local, 00 UTC.	49
4.5	Radiossondagens com lançamento à tarde. (a) Radiossondagem na cidade de Peachtree, às 18 : 30 h no horário local, 00 UTC. (b) Radiossondagem na cidade de Amarillo, às 17 : 30 h no horário local, 00 UTC.	50
4.6	Lançamento às 3:30 h para a estação de Reno, fevereiro.	51
4.7	Lançamento às 5:30 h para a estação de Amarillo.	52
4.8	Lançamento às 6:30 h para a estação de Peachtree.	52
4.9	Lançamento às 6:30 h para a estação de Peachtree.	53
4.10	Lançamento às 05:30 h, horário local, para a estação de Amarillo.	54
4.11	Lançamento às 15:30 h, horário local, para a estação de Reno.	55
4.12	Gráfico de dispersão para a configuração st4.	56
4.13	Gráfico de dispersão para a configuração sth10.	57
4.14	Gráfico de dispersão para a configuração sta8.	58
4.15	Gráfico de dispersão para a configuração sta5.	59

Lista de Tabelas

2.1	Erros obtidos pelo modelo de crescimento em Tijucas do Sul.	24
2.2	Erros obtidos pelo modelo de crescimento no Kansas.	24
3.1	Dados retirados das radiossondagens de alta resolução.	26
3.2	Identificação das estações do ano para entrada na RNA	35
3.3	Diferentes configurações de entrada para treinamento e validação da RNA.	37
3.3	Diferentes configurações de entrada para treinamento e validação da RNA.	38
4.1	Resultados obtidos para z_0 para cada localidade.	41
4.2	Resultados obtidos com emprego da parametrização de Panofsky.	42
4.3	Resultados obtidos para as melhores configurações do bloco de experimen- tos <i>st</i>	55
4.4	Resultados obtidos para as piores configurações do bloco de experimentos <i>st</i> .	56
4.5	Resultados obtidos para as melhores configurações do bloco de experimen- tos <i>sth</i>	57
4.6	Resultados obtidos para as melhores configurações do bloco de experimen- tos <i>sta</i>	58
4.7	Resultados obtidos com a separação dos dados do dia e da noite.	59
4.8	Resultados obtidos para as melhores configurações com emprego da RNA generalizada.	62
A.1	Parâmetros de qualidade para as configurações testadas	67

Lista de Símbolos

η	taxa de aprendizado da rede
Γ_d	lapse rate em condições secas
κ	constante de vón Kármán
Ψ_m, Ψ_h	funções de estabilidade
ρ	densidade do ar
$\sigma(x)$	função de ativação da rede
σ_x^2	desvio padrão de uma variável x
τ	tensão de cisalhamento
θ	temperatura potencial
θ_*	fluxo de calor superficial
θ_0	temperatura no nível da superfície
θ_v	temperatura potencial virtual
ζ	parâmetro de estabilidade de Monin-Obukhov
b	bias
c_p	calor específico a pressão constante
E_i	valor estimado
est	estação/período do ano
f	parâmetro de Coriolis
g	aceleração da gravidade

G_0	velocidade do vento geostrófico de superfície
H	fluxo de calor sensível
h	altura da Camada Limite Atmosférica
H_v	fluxo de calor sensível virtual
h_{g500}	altura geopotencial do nível de 500 hpa
h_{g850}	altura geopotencial do nível de 800 hpa
h_{gpt}	altura geopotencial
h_{old}	última altura da CLA medida
$hora$	hora de lançamento da radiossonda
IOA	Index of Agreement
IOA_{mod}	Index of Agreement modificado
L	comprimento de Monin-Obukhov
$lapse$	distância temporal o momento de estimativa e a última medição de h
lat	latitude
MAE	Média dos Valores Absolutos dos Erros
MBE	Média dos Valores dos Erros
N	refratividade
O_i	valor observado
P_0	pressão no nível da superfície
q	umidade específica
r	Correlação Linear
Ri_b	número de Richardson para uma camada
Ri_{bs}	número de Richardson para uma camada suavizado por média móvel
$RMSE$	Raiz do Erro Médio Quadrático

T	Temperatura
T_0	temperatura no nível da superfície
u_*	velocidade de atrito
U_1	componente da velocidade horizontal do vento na direção x
U_2	componente da velocidade horizontal do vento na direção y
u_3	componente da velocidade do vento na direção z
U_H	velocidade média do vento, na horizontal
u_r	velocidade horizontal em um nível de referência z_r
U_{1-500}	componente da velocidade do vento a 500 hpa na direção x
U_{1-850}	componente da velocidade do vento a 850 hpa na direção x
U_{2-500}	componente da velocidade do vento a 500 hpa na direção y
U_{2-850}	componente da velocidade do vento a 850 hpa na direção y
U_{H0}	velocidade horizontal do vento no nível da superfície
UR	umidade relativa
UR_0	umidade relativa no nível da superfície
UR_{700}	umidade relativa no nível de 700 hpa
v	velocidade média constante de um escoamento
w_i	peso sináptico
X_i	parte determinística de x_i
x_i	variável de interesse
x'_i	flutuação turbulenta de x_i
z	altura em relação ao solo
z_0	altura de rugosidade
z_f	altura final em relação ao solo

z_r	nível de referência
z_{0r}	valor intermediário para a obtenção de z_0
z_{in}	altura inicial em relação ao solo

Sumário

1	Introdução	1
1.1	Objetivos	2
1.1.1	Objetivo geral	2
1.1.2	Objetivos específicos	2
1.2	Estrutura do Trabalho	3
2	Revisão bibliográfica e fundamentação teórica	4
2.1	Camada Limite Atmosférica	4
2.1.1	Turbulência	5
2.1.2	Estabilidade na Camada Limite Atmosférica	6
2.2	Comprimento de estabilidade de Monin-Obukhov	7
2.2.1	Altura da camada limite atmosférica	9
2.2.2	Radiossondagens	11
2.3	Redes Neurais Artificiais	12
2.3.1	<i>Perceptron</i> multicamadas	13
2.3.2	Treinamento	15
2.4	Estimativa da altura da CLA	17
2.4.1	Métodos de estimativa com dados de perfis de variáveis atmosféricas	18
2.4.2	Métodos de estimativa a partir de dados de superfície	21
3	Materiais e Métodos	25
3.1	Base de dados	25
3.1.1	Dados de radiossondagens	25
3.1.2	Procedimentos para análise	26
3.1.3	Caracterização dos sítios analisados	28
3.1.4	Estimativa da altura da CLA a partir de radiossondagens	32
3.2	Parametrização da altura da CLA	32
3.3	Treinamento e validação da Rede Neural Artificial	33
3.4	Parâmetros para a avaliação de qualidade das estimativas	39

4	Resultados e discussão	41
4.1	Altura de rugosidade	41
4.2	Estimativa da altura da CLA pela parametrização de Panofsky	42
4.3	Estimativa da altura da CLA a partir de radiossondagens	45
4.3.1	Camada de mistura - Inverno	45
4.3.2	Camada de mistura - Verão	48
4.3.3	Camada noturna	48
4.3.4	Dificuldades técnicas	51
4.4	Estimativa da altura da CLA com emprego da RNA	53
4.4.1	Estimativa de h com emprego de uma RNA generalizada	54
4.4.2	Estimativa de h com emprego de duas RNA's especializadas	58
4.5	Discussão	59
5	Conclusão	63
A	Parâmetros estatísticos	67

Capítulo 1

Introdução

A atmosfera é um envoltório gasoso ao redor do nosso planeta, e influencia todos os momentos de nossas vidas. Seu limite inferior é a superfície da terra, e é na interface superfície-atmosfera que ocorrem trocas importantes de diversas grandezas físicas, como umidade e energia. A influência da superfície sobre a atmosfera produz o comportamento diferenciado observado em sua porção inferior. Assim, para estudarmos esse efeito, dividimos a troposfera em duas camadas: a parte superior, denominada atmosfera livre e palco de fenômenos meteorológicos de grande escala, e a parte inferior, com intensidade de turbulência significativamente maior, denominada Camada Limite Planetária, ou Camada Limite Atmosférica (CLA).

Passamos nossas vidas, ou a maior parte delas, na superfície terrestre, imersos no ar da Camada Limite Atmosférica, submetidos aos fenômenos que nela ocorrem, sendo muitas vezes dependentes deles. A troca de umidade entre a superfície e a atmosfera é um bom exemplo da importância da dinâmica da CLA em nossas vidas. A atmosfera livre pode levar uma partícula ao longo de quilômetros de distância por advecção, mas é o escoamento turbulento característico da CLA que levará essa partícula verticalmente, partindo da superfície e alcançando o topo da camada, que usualmente varia de 100 m a 4000 m. Numa visão bastante simplificada, a umidade proveniente da superfície é elevada ao nível de condensação por levantamento, iniciando o processo de formação de nuvens e precipitação posterior, fazendo da dinâmica da CLA um processo essencial ao ciclo hidrológico.

Da mesma forma que o vapor d'água, outros gases e material particulado lançados na atmosfera, seja por carros, pequenas chaminés, grandes indústrias, ou mesmo gases liberados em grandes lagos, são misturados e carregados a grandes distâncias de seu lançamento, num processo denominado dispersão, que só acontece dessa maneira pela existência da turbulência. Contudo, ao alcançar o topo da CLA encontramos em algumas ocasiões uma espécie de “tampa”, a qual retém grande parte de todos esses gases e poluentes, limitando

a parcela da atmosfera na qual teremos a dispersão de poluentes. Nesses processos, a CLA é nosso volume de controle, e conhecer sua dimensão vertical é essencial para a aplicação em modelos de dispersão. Contudo, essa dimensão, conhecida como altura da Camada Limite Atmosférica (H), varia ao longo do tempo e do espaço, com relações de formação e desenvolvimento apenas parcialmente conhecidas.

Sabemos que com o aquecimento da superfície pela radiação solar temos uma intensificação da turbulência na camada, e seu crescimento ao longo do dia. Da mesma forma, o resfriamento noturno promove uma mudança em sua estrutura, a partir de sua base na superfície. A distribuição vertical da velocidade do vento também está relacionada às características do escoamento na camada, e conseqüentemente à sua extensão vertical. Dessa forma, um grande número de variáveis, possivelmente relevantes, foi elencado e estudado ao longo dos anos, na tentativa de obter formulações para determinação de H a partir de parâmetros conhecidos, mas, de forma geral, os resultados não foram satisfatórios [1].

Assim, surgem diferentes métodos de estimativa desse parâmetro a partir de perfis de escalares, como a aplicação de dados obtidos por SODAR [1] e LIDAR [2], e modelos para obtenção de H a partir de perfis provenientes de radiossondagens, de forma automatizada [3]. Os métodos representam avanços significativos, mas ainda há limitações, e dificuldades operacionais, como custo de lançamento de radiossondas, ou o alcance vertical dos equipamentos SODAR e LIDAR.

Isto posto, a busca por novas metodologias para estimativa de H continua, e desperta o interesse da aplicação de métodos de inteligência artificial, dada sua capacidade de processamento de problemas não-lineares e complexos, categoria de problemas na qual a estimativa de H se encaixa perfeitamente.

Nesse trabalho, pretendemos avaliar a possibilidade de determinação da altura da CLA com uso de dados de superfície em estações meteorológicas.

1.1 Objetivos

1.1.1 Objetivo geral

Este trabalho tem como objetivo geral avaliar o desempenho de uma rede neural artificial do tipo *multilayer perceptron* na estimativa da altura da camada limite atmosférica, utilizando dados de superfície como entrada.

1.1.2 Objetivos específicos

Com o intuito de desenvolver um método de estimativa da altura da camada limite, faz-se necessária a ampliação do conhecimento sobre a mesma. Isto posto, torna-se claro

o objetivo de aprofundar-se na análise dos parâmetros considerados pela literatura como relevantes ao objetivo geral.

Para aplicação de uma rede neural artificial, uma base de dados confiável e representativa é essencial, e dados de H_{CLA} medidos não estão disponíveis em bases de dados existentes, sendo necessária a obtenção desses dados especificamente para esse trabalho.

Metodologias existentes para estimar a altura da camada limite atmosférica serão estudadas e implementadas, de forma a avaliar qual a melhor estratégia para aplicação desses métodos e formação da base de dados com informação de H_{CLA} .

Além disso, o desempenho de uma rede neural artificial única para estimativa de H_{CLA} para casos instáveis e estáveis será comparado ao desempenho de redes especializadas para cada uma dessas situações.

1.2 Estrutura do Trabalho

Este trabalho é estruturado da seguinte maneira. No Capítulo 2 são apresentados conceitos necessários à compreensão do trabalho, bem como uma revisão bibliográfica com um pouco do que se encontra disponível sobre o tema na literatura científica. No Capítulo 3 são descritos os materiais utilizados, como a fonte de dados e o programa da Rede Neural Artificial (RNA), e também a metodologia pela qual foi desenvolvida a base de dados utilizada, e sua aplicação para treinamento e validação da RNA. O Capítulo 4 contém os resultados encontrados ao longo do trabalho, sendo eles a avaliação da parametrização de PANOFSKY para camada limite instável, a base de dados desenvolvida a partir de perfis verticais de temperatura potencial virtual e de umidade específica, e a avaliação do desempenho do emprego da RNA para estimativa da altura da camada limite atmosférica. Por fim, o Capítulo 5 apresenta as principais conclusões sobre o trabalho desenvolvido e os resultados obtidos.

Capítulo 2

Revisão bibliográfica e fundamentação teórica

2.1 Camada Limite Atmosférica

Para compreender o comportamento da camada limite atmosférica, vamos primeiro abordar brevemente a formação de camadas-limite em escoamentos sobre paredes. Segundo TENNEKEES e LUMLEY [4], por menor que seja a viscosidade de um fluido escoando sobre uma superfície, teremos imposta a condição de contorno de não-deslizamento, sendo então a velocidade do fluido na superfície sólida igual à velocidade da própria superfície.

Consideramos então um escoamento sobre uma superfície plana, com velocidade média constante v . Ao nos aproximarmos da superfície, podemos observar uma região com velocidade média variável, decaindo até satisfazer a condição de não deslizamento. Nessa região, denominada camada limite, o escoamento sofre interferência direta da superfície sobre a qual escoa, e podemos notar variações abruptas em propriedades do fluido como temperatura, velocidade e densidade [5].

Analogamente ao apresentado, podemos considerar a superfície terrestre como sendo a parede sobre a qual nosso fluido escoa. O fluido em questão é o próprio ar atmosférico, que está em constante movimento. A camada limite atmosférica (CLA), ou camada limite planetária, é definida na literatura, de forma usual, como sendo a porção inferior da atmosfera, na qual podem ser sentidos os efeitos diretos da interação com a superfície [6]. GARRATT [7], de forma mais detalhada, a define também como a porção inferior da atmosfera sob influência direta da superfície, mas com os efeitos sofridos em uma escala menor que um dia, e na qual os fluxos de calor e quantidade de movimento, por exemplo, se dão por meio de vórtices turbulentos numa escala da mesma ordem de grandeza da profundidade da camada ou menor. Para STULL [8], a escala de tempo é de uma hora

ou menos.

Propriedades como umidade e temperatura, e os efeitos do atrito com a superfície (seja ela continente ou oceano) são rapidamente transmitidos ao longo da CLA, por meio dos vórtices turbulentos nela presentes. A natureza turbulenta da CLA talvez seja a mais importante de suas características, sendo muitas vezes utilizada inclusive para defini-la, e será, portanto, mais detalhada em uma seção seguinte.

2.1.1 Turbulência

A maior parte dos escoamentos observados na natureza são turbulentos. Entretanto, a compreensão do fenômeno da turbulência em si requer um certo grau de abstração por parte de quem se propõe a tal. Uma visão interessante e de caráter bastante didático é a de Richardson, na qual a turbulência se apresenta na forma de turbilhões, ou vórtices, de diferentes tamanhos, num processo de “alimentação” dos menores pelos maiores, até alcançar a escala da dissipação viscosa [9].

Uma saída para contornar a dificuldade de descrever o fenômeno em si é detalhar e explicar suas características mais relevantes, sendo essa também a linha a ser seguida neste trabalho.

Tomando a visão de Richardson, podemos enxergar a turbulência, ou ao menos tentar fazê-lo, como uma infinidade de vórtices de diferentes tamanhos. A energia é transferida do escoamento médio para os vórtices turbulentos a partir das maiores escalas, ou maiores vórtices, enquanto que nas menores escalas há dissipação por meio da viscosidade, devido ao atrito, formando a então denominada cascata de energia [5].

A cascata de energia é uma importante característica dos escoamentos turbulentos, transferindo continuamente a energia cinética do escoamento médio para as menores escalas, sendo então transformada em energia interna e calor, através da viscosidade. Portanto, a turbulência precisa de um fornecimento constante de energia para sua manutenção, decaindo rapidamente sem ele [4, 5].

Outra característica de extrema importância para os escoamentos turbulentos é sua difusividade, inúmeras ordens de grandeza maior que a difusividade molecular [5]. Para esse autor, essa é a mais importante das características da turbulência, e a eficiência na difusão de quantidade de movimento, calor e massa nos escoamentos turbulentos deve-se a ela. No escopo da CLA, a dispersão de poluentes e evaporação são em grande parte devidas à difusividade da turbulência.

Flutuações aleatórias na vorticidade são características de escoamentos turbulentos, e não poderiam ser mantidas apenas com flutuações bidimensionais de velocidade [4]. Isto posto, observamos que o campo de velocidade para qualquer escoamento turbulento (salvo exceções englobadas por escoamentos atmosféricos de grande escala) é tridimensional, bem

como o campo de vorticidade, e altamente variável no tempo e espaço [5].

Começamos a perceber então uma característica essencial para explicar a abordagem adotada pelos estudiosos ao tratar de escoamentos turbulentos: a natureza da turbulência como sendo estocástica. Simulações numéricas baseadas nas equações de Navier-Stokes mostraram-se extremamente sensíveis, mesmo a pequenas variações nas condições iniciais e de contorno, tornando uma abordagem estatística muito mais interessante [5]. Segundo DIAS [9], a aplicação da teoria de probabilidade e processos estocásticos faz-se essencial para o estudo da turbulência.

Ao observarmos medições de velocidade do vento na CLA, ao longo de um período contínuo, podemos identificar a velocidade média do escoamento, com variações em grande escala, e flutuações instantâneas em relação a essa média (negativas e positivas), chamadas de flutuações turbulentas.

Da mesma forma, outras variáveis de interesse no escoamento turbulento podem ser representadas por suas médias e flutuações turbulentas, utilizando a decomposição de Reynolds. Assim, uma variável x_i pode ser separada entre sua parte determinística X_i , e sua flutuação turbulenta x'_i [9]:

$$x_i = X_i + x'_i \quad (2.1)$$

2.1.2 Estabilidade na Camada Limite Atmosférica

Considerando uma abordagem genérica, podemos descrever a estabilidade de um sistema como a resposta do mesmo frente à exposição a perturbações. A amplificação ou amortecimento das perturbações caracterizam, respectivamente, sistemas instáveis e estáveis. Em um sistema neutro não há tais tendências de resposta observadas [10].

A estabilidade do escoamento está fortemente relacionada à presença de turbulência, bem como a sua intensidade. Escoamentos estáveis são ou passam a ser laminares, enquanto que escoamentos instáveis tendem a tornar-se ou permanecerem turbulentos [8]. Tal condição é de extrema importância para conhecermos a dinâmica do escoamento em estudo. Na CLA, as condições de estabilidade alteram a capacidade de mistura de escalares de interesse, e nos dizem muito sobre o que esperar quanto à espessura da camada.

Quando avaliada estaticamente, a estabilidade representa, basicamente, estabilidade termodinâmica, e se refere à flutuabilidade das parcelas de ar. Nesse tipo de consideração não se levam em conta os efeitos mecânicos, mas apenas a diferença de densidade entre porções de ar e o ambiente, gerando movimentos convectivos ou subsidência.

Se tomarmos uma parcela em um nível inicial z_{in} e a movimentarmos verticalmente a um nível $z = z_{in} + \Delta z$, podemos ter três diferentes cenários para a posição final z_f . Em um

ambiente instável, a parcela tende a continuar o movimento, amplificando a perturbação. Já em um ambiente estável, a perturbação é amortecida, e $z_f = z_{in}$. No terceiro cenário a perturbação não é amplificada, tampouco é amortecida, o ambiente é considerado neutro e $z_f = z$.

Contudo, mesmo em uma camada considerada estaticamente estável pode haver instabilização de origem mecânica, devido ao cisalhamento do vento. Assim, uma maneira alternativa de avaliação é a estabilidade dinâmica, na qual são consideradas tanto forças mecânicas quanto térmicas.

Em aplicações meteorológicas, é comum o uso do número de Richardson para uma camada (Ri_b)

$$Ri_b = \frac{|g| \Delta \bar{\theta}_v \Delta z}{\bar{\theta}_v \left[(\Delta \bar{U}_1)^2 + (\Delta \bar{U}_2)^2 \right]} \quad (2.2)$$

como parâmetro para determinar a estabilidade. Em 2.2 g é a aceleração da gravidade, $\Delta \bar{\theta}_v$ é a variação vertical da temperatura potencial virtual média, $\bar{\theta}_v$ é a temperatura potencial virtual média na camada, e $\Delta \bar{U}_1$ e $\Delta \bar{U}_2$ são as variações dos valores das componentes horizontais do vento U_1 e U_2 .

2.2 Comprimento de estabilidade de Monin-Obukhov

A escala de comprimento

$$L = \frac{u_*^3}{\kappa g \frac{-H_v}{c_p \rho T}} \quad (2.3)$$

foi primeiramente apresentada por Alexander Obukhov em 1946, e sua discussão pode ser encontrada em [11]. Na Equação (2.3), vemos que L carrega informações sobre o escoamento, sua interação com a superfície, e transferência de calor, pois u_* é a velocidade de atrito definida como $\sqrt{\tau/\rho}$, sendo τ a tensão de cisalhamento, κ é a constante de *vón Kármán* com valor de 0.4, H_v é o fluxo de calor sensível virtual, c_p é calor específico a pressão constante, ρ é densidade do ar e T é a temperatura absoluta.

A validade de L abrange a camada superficial, na qual é considerada uniformidade de propriedades horizontais, tal qual um escoamento sobre uma placa plana rugosa de dimensões infinitas. As características do escoamento médio são, portanto, funções apenas de z .

O trabalho apresentado por Obukhov foi posteriormente desenvolvido por Monin, e em 1956 foi publicado um artigo de autoria conjunta dos dois pesquisadores russos, trazendo o conceito de z_0 baseado no perfil de velocidade logarítmico. Dependente apenas das

características de cobertura da superfície, z_0 pode ser conhecido a partir de observações do perfil de vento, utilizando

$$u_H(z) = \frac{u_*}{\kappa} \ln \left(\frac{z}{z_0} \right), \quad (2.4)$$

na qual u_H é a velocidade média horizontal, e z é a altura, tendo validade para condições próximas à neutralidade.

A tensão de cisalhamento é considerada então uma tensão de cisalhamento turbulenta, descrita por

$$\tau = -\overline{\rho u'_H u'_3}, \quad (2.5)$$

e u'_H e u'_3 são flutuações das componentes horizontal e vertical de velocidade, respectivamente. O fluxo de calor também passa a ser apresentado em função de flutuações,

$$H_v = c_p \overline{\rho u'_3 T'}. \quad (2.6)$$

Os fluxos turbulentos de temperatura e quantidade de movimento são considerados constantes para a camada superficial, e apenas para essa região onde tais considerações podem ser feitas tem validade a teoria de similaridade de Monin-Obukhov.

Com uma manipulação dos parâmetros, pode ser encontrada uma representação de L em função do fluxo de calor na superfície θ_* , como mostra a Equação (2.7), apresentada por ARYA [12]. A razão entre a altura z e o comprimento de Obukhov L é o parâmetro de similaridade ou estabilidade de Monin-Obukhov, adimensional, representado por ζ . Em [12] também são apresentadas as equações normalizadas para os perfis de velocidade e temperatura potencial em função de z , com as quais podemos obter

$$L = \frac{u_*^2}{\kappa(g/T_0)\theta_*}, \quad (2.7)$$

$$u_* = \frac{\kappa \overline{u_r}}{[\ln(z_r/z_0) - \Psi_m(z_r/L)]}, \quad (2.8)$$

e

$$\theta_* = \frac{\kappa \Delta \bar{\theta}}{\alpha [\ln(z_2/z_1) - \Psi_h(z_2/L) + \Psi_h(z_1/L)]}, \quad (2.9)$$

para determinação de u_* e θ_* . Nas Equações 2.7, 2.8 e 2.9, T_0 é a temperatura do ar a um estado de referência, z_r é a altura de medição da velocidade de referência $\overline{u_r}$, e z_1 e z_2 são duas alturas de medição de θ para obtenção de $\Delta\theta$.

Em (2.8) e (2.9), Ψ_m e Ψ_h são funções de estabilidade, cujas formulações aproximadas

determinadas empiricamente são

$$\Psi_m = \Psi_h = -\frac{5z}{L}, \quad \text{para} \quad \frac{z}{L} \geq 0, \quad (2.10)$$

$$\Psi_m = 2 \ln \left(\frac{1+x}{2} \right) + \ln \left(\frac{1+x^2}{2} \right) - 2 \arctg(x) + \frac{\pi}{2} \quad \text{para} \quad \frac{z}{L} < 0, \quad (2.11)$$

e

$$\Psi_h = 2 \ln \left(\frac{1+x^2}{2} \right), \quad \text{para} \quad \frac{z}{L} < 0, \quad (2.12)$$

sendo $x = (1 - 15\zeta)^{1/4}$. A teoria de Monin-Obukhov é limitada pelas condições de estabilidade, sendo considerada válida dentro dos limites de $-5 < \zeta < 2$ [5].

2.2.1 Altura da camada limite atmosférica

A Camada Limite Atmosférica não possui uma delimitação exata, mas sim uma faixa de transição, de um ambiente com fluxos turbulentos significativos para a atmosfera livre. Em dias ensolarados e quentes, a produção de turbulência é grande e a importância desses fluxos turbulentos pode se estender a altos níveis na atmosfera, chegando a 2 ou 3 quilômetros. Em contrapartida, ao analisarmos um cenário noturno, por exemplo, há grandes chances de percebermos um *lapse rate* (taxa de resfriamento na vertical) estável imposto pelo resfriamento do solo, amortecendo a produção da turbulência e resultando em uma CLA mais fina e bem comportada. [6].

Torna-se, então, mais clara a dificuldade de estabelecer um topo para a CLA, podendo esta ser definida como a altura da base de uma inversão elevada de temperatura potencial [1], em casos de camadas bem misturadas e desenvolvidas, ou como a altura do mais baixo máximo no perfil de velocidade do vento, conhecido como jato de baixa altitude, para CLA estáveis [13].

Dado que o topo da CLA é marcado por mudanças abruptas em perfis de propriedades escalares, a observação da umidade relativa e da umidade específica é de extrema importância.

Assim, são apresentados alguns métodos para determinação da altura da CLA (h) a partir de perfis como os fornecidos por radiossondagens, encontrados em [14].

1. Método da parcela (*parcel method*), o qual considera o nível no qual uma parcela originária da superfície e em processo de levantamento encontraria o equilíbrio com o ambiente. Tal comportamento seria observado a uma altura na qual θ_v do ambiente seja o mesmo da superfície.

2. A base de uma inversão elevada de θ_v , como mencionado previamente, funciona como uma tampa, retendo a mistura de escalares para níveis superiores, sendo também uma maneira de identificar h .
3. Nível do máximo gradiente vertical de θ_v , em muitos casos coincidente com a inversão elevada.
4. Nível do topo de uma inversão com base na superfície, que se aplica a camadas estáveis.
5. Nível do mínimo gradiente vertical de umidade relativa (UR), identificando a variação abrupta de propriedades na interface CLA - Atmosfera livre. Pode ser utilizado também o mínimo gradiente de umidade específica (q).
6. Nível do mínimo gradiente de refratividade (N).
7. Número de Richardson volumétrico ($Ri_{bc} \geq 0.25$).

Além de métodos baseados em perfis, são propostas parametrizações na literatura, em sua maior parte para camadas estáveis. Em 1972, por exemplo, ZILITINKEVICH [15] propôs a relação

$$h = d \left(L \frac{u_*}{f} \right)^{1/2}, \quad (2.13)$$

na qual d é uma constante de ordem um e f é o parâmetro de Coriolis. Também há para casos estáveis a proposição de Laikhtman,

$$h = cG_0 \left(\frac{g}{T_0} \left\langle \frac{\partial \bar{\theta}}{\partial z} \right\rangle \right)^{-1/2}, \quad (2.14)$$

considerada num estudo de 1969 como a melhor das formulações propostas até então para a estimativa da altura da camada limite em condições estáveis. Nessa parametrização, G_0 é a velocidade do vento geostrófico de superfície e c é uma constante [16].

Para condições instáveis, podemos encontrar em [17] a parametrização proposta por PANOFISKY,

$$h = \frac{T_h - T_{sunrise}}{\Gamma_d - \Lambda}, \quad (2.15)$$

na qual Γ_d é o *lapse rate* em condições secas, e Λ é o *lapse rate* no momento do nascer do sol. $T_{sunrise}$ é a temperatura medida no momento do nascer do sol, e T_h a temperatura no momento de estimativa.

2.2.2 Radiossondagens

Sondagens atmosféricas são medições de perfis verticais de parâmetros atmosféricos relevantes, como temperatura, umidade, pressão e velocidade do vento [18]. Em radiossondagens, os sensores são usualmente carregados verticalmente em ascensão livre por um balão atmosférico, e os dados são transmitidos em tempo real para um receptor em solo, a uma frequência que pode variar de sensor para sensor. Radiossondas de alta resolução enviam dados a uma frequência de 1 s, enquanto radiossondas convencionais o fazem a uma frequência de 6 s.

O instrumento de sondagem consiste em uma espécie de caixa, como a mostrada na Figura 2.1a, contendo os sensores de umidade, temperatura e pressão, e também um GPS, para obtenção da velocidade horizontal do vento. Essa caixa é presa a um balão atmosférico feito de látex (Figura 2.1), que é inflado a um diâmetro de aproximadamente 1.8 m (6 pés). O balão ascende e pode chegar a 30000 metros de altitude, quando atinge um diâmetro muito maior que o de seu lançamento, aproximadamente 8 m, e então estoura. O aumento de volume e consequente estouro se deve a queda de pressão ao longo da subida na atmosfera, que leva o ar dentro do balão a se expandir.



(a)



(b)

Figura 2.1: Imagem de uma radiossonda com receptor GPS acoplado (a) e de um balão atmosférico desinflado (b).

Fonte: NOAA

A sonda é equipada com um paraquedas, visando evitar possíveis danos e acidentes que poderiam ser causados por uma queda do equipamento em alta velocidade.

2.3 Redes Neurais Artificiais

Relações entre variáveis na natureza são quase sempre bastante complicadas e altamente não-lineares. Hidrologia, interações ecológicas, climatologia, oceanografia, micrometeorologia, dentre outras áreas, apresentam dificuldades de análise devido a sua complexidade. A observação de um sistema tão complexo quanto os problemas a serem solucionados, o cérebro humano, levou ao desenvolvimento de algoritmos de inteligência artificial, os quais tentam simular propriedades que fazem desse órgão uma ferramenta tão poderosa, sendo os principais: não-linearidade, paralelismo e generalização.

O neurônio biológico, representado na Figura 2.2, pode receber sinais de diversos outros neurônios simultaneamente, através dos dendritos (conexões de entrada), os quais são processados no soma (corpo celular) e transferidos através do axônio (conexões de saída) para outros neurônios. Esse processo é denominado sinapse, e pode ser químico ou elétrico, sendo o químico mais comum. Em ambos os casos, os diversos estímulos são “contabilizados” no soma, e apenas atingindo um determinado valor de polarização teremos a continuidade da transmissão do impulso. Do contrário, o processo se encerra e não há a transferência de sinais para neurônios subsequentes. A diferença principal entre os dois tipos de sinapse é a unidirecionalidade das sinapses químicas, diferentemente das elétricas que podem ocorrer em ambos os sentidos.

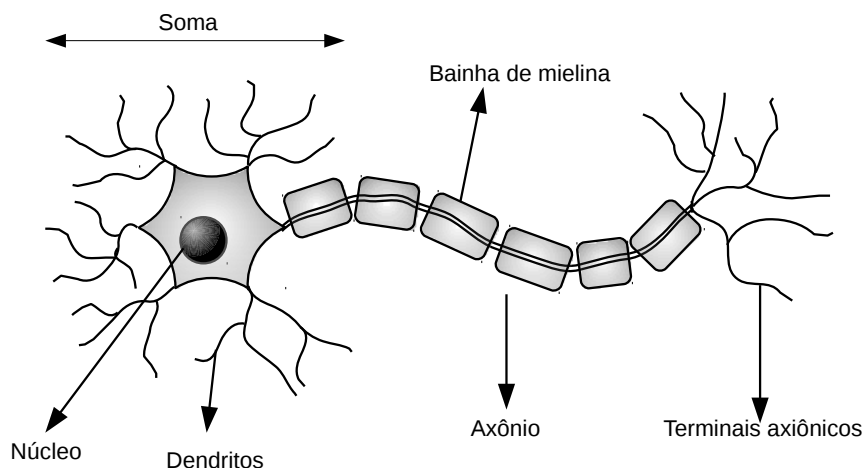


Figura 2.2: Representação de um neurônio biológico.

Da mesma forma que o neurônio biológico, o neurônio artificial, representado esquematicamente na Figura 2.3, pode receber diversos estímulos, representados pelos n dados

de entrada x_i do modelo.

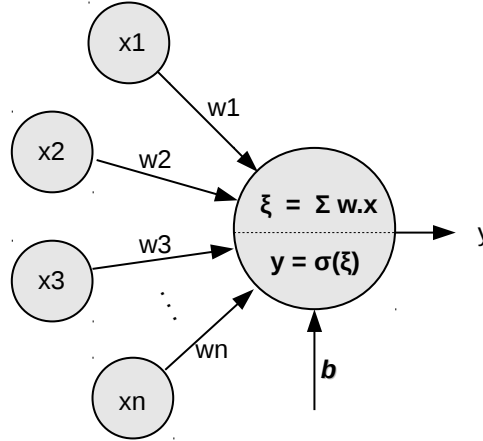


Figura 2.3: Representação de um neurônio artificial.

A cada um desses dados é aplicado um peso w_i , e uma função de ativação σ , como exemplificado em

$$\begin{cases} \xi = \sum_{i=1}^n x_i w_i \\ y = \sigma(\xi + b) \end{cases}, \quad (2.16)$$

é então aplicada ao somatório das n entradas, determinando se o neurônio será ou não ativado. Na Equação (2.16), b representa o *bias*, um viés atuando como valor limitante adicional.

2.3.1 *Perceptron* multicamadas

O primeiro modelo de RNA foi apresentado por Rosenblatt em 1958, e consistia na combinação de unidades de processamento, os neurônios artificiais. Nesse modelo, denominado *perceptron*, foi apresentada a ideia de separação de neurônios com mesma função. Os neurônios da camada de entrada recebem informações, e os neurônios da camada de saída as processam e expõem, sendo que cada neurônio da camada de entrada está ligado a todos os neurônios da camada de saída, mas não há conexões entre neurônios de uma mesma camada, como mostra a Figura 2.4. O fluxo de informações acontece apenas em um sentido, como nas sinapses químicas do cérebro humano, caracterizando uma rede do tipo *feedforward*. Com o passar do tempo, foram desenvolvidas também redes com retroalimentação, nas quais os dados de saída são reinseridos na rede como entrada.

Contudo, a existência de apenas duas camadas impõe limitações ao modelo de Rosen-

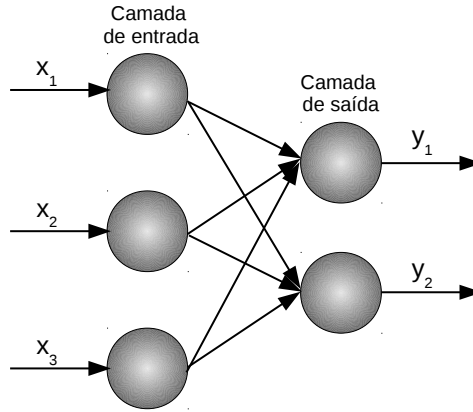


Figura 2.4: Representação de uma rede neural artificial do tipo *perceptron*.

blatt, e a rede *perceptron* é capaz de resolver apenas problemas de classificação linearmente separáveis. Tal limitação levou ao desenvolvimento de um *perceptron* com múltiplas camadas colocadas entre a camada de entrada e a de saída, o *perceptron* multicamadas ou *multilayer perceptron*. Os neurônios dessas camadas intermediárias, também denominadas ocultas ou internas, são destinados apenas ao processamento de dados, e as camadas de entrada e de saída continuam com as funções de recebimento de informações e exposição de resultados. Na Figura 2.5 vemos uma representação de uma rede *perceptron* multicamadas com uma camada oculta. Assim como no modelo *perceptron* de uma camada, são aplicados pesos às conexões entre dois neurônios de camadas consecutivas, e uma conexão entre o neurônio i da camada $l - 1$ e o neurônio j da camada l tem peso w_{ij}^l . Em cada neurônio é realizado o somatório dos *inputs*, e a esse somatório é então aplicada uma função de ativação, conforme apresentado a seguir, numa variação da Equação (2.16).

$$\begin{cases} \xi_i^l = \sum_{i=1}^n x_i^{l-1} w_{ij}^l \\ y = x_i^l = \sigma(\xi_i^l + b) \end{cases} \quad (2.17)$$

O número de camadas ocultas, quantos neurônios devem ser colocados em cada uma delas e as ligações entre os neurônios caracterizam a arquitetura da rede, sendo usualmente definidos empiricamente. Já a função de ativação utilizada em cada camada depende do tipo de problema, de como tratamos os dados de entrada e também do papel que essa camada desempenha na rede.

Há várias formas disponíveis na literatura para a função de ativação σ , dentre as quais podemos destacar a função linear, a função degrau e as funções do tipo sigmóide. O objetivo da aplicação de uma função de ativação degrau ou sigmóide é limitar a amplitude dos dados de saída da RNA [19]. No caso da camada de saída, cujos resultados podem assumir quaisquer valores, é usualmente empregada a função linear. Já nas camadas

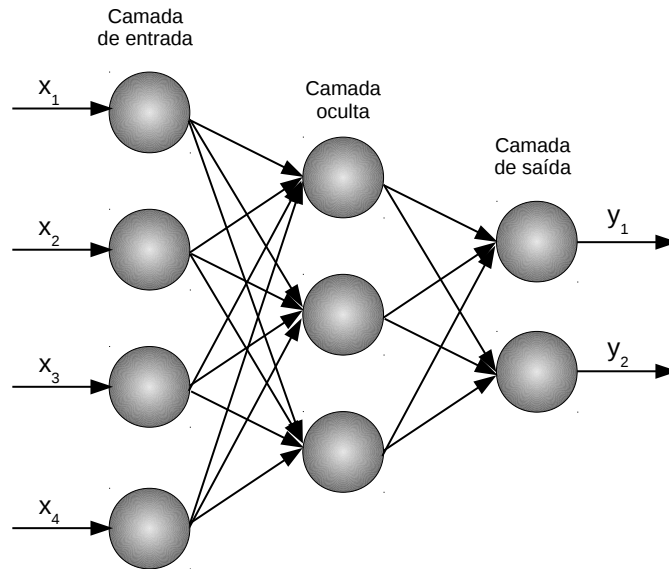


Figura 2.5: Representação de uma rede neural artificial do tipo *perceptron* multicamadas.

internas de redes do tipo multicamadas, a aplicação de funções lineares não possibilita a capacidade de processamento de problemas não-lineares, e é necessária a aplicação de funções que representem a não-linearidade do problema, como funções do tipo sigmóide. Uma função bastante usada para redes com dados de entrada do tipo binário, ou seja, no intervalo $[-1,1]$ é a função tangente hiperbólica,

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}. \quad (2.18)$$

2.3.2 Treinamento

A capacidade de aprendizado é um componente essencial a um sistema de inteligência artificial. Para desempenhar uma determinada função, assim como no caso de um cérebro biológico, a rede neural deverá antes ser treinada para aquele problema. Há três maneiras de realizar o treinamento de uma rede neural: supervisionado, não supervisionado e por reforço.

No aprendizado supervisionado, um agente externo à rede apresenta um padrão de referência, ou seja, a cada entrada e processamento de dados, o resultado obtido pela rede com seu vetor de pesos atual será comparado com o resultado correto. O aprendizado por reforço também conta com um agente externo, mas em lugar da apresentação do resultado

esperado, o resultado fornecido pela rede recebe uma nota de acordo com sua acurácia. Diferentemente dos dois anteriores, no aprendizado não supervisionado não são fornecidas referências para a rede.

Um algoritmo bastante comum para redes *perceptron* multicamadas é o *backpropagation*, um método supervisionado de ajustes de pesos sinápticos por correção de erros. Ao iniciar o treinamento, o vetor de pesos é carregado com valores aleatórios, que serão ajustados conforme os dados de entrada sejam apresentados e processados pela rede, e o resultado do processamento seja comparado ao resultado esperado. O erro é calculado numa determinada iteração n de acordo com

$$e_k(n) = d_k(n) - x_k(n), \quad (2.19)$$

e é a diferença entre o valor esperado $d_k(n)$ e o valor obtido $x_k(n)$. O erro calculado será aplicado para correção do peso w_{ik} pela *regra delta*,

$$\Delta w_{ik}(n) = \eta e_k(n) x_i(n), \quad (2.20)$$

onde η é a taxa de aprendizado da rede, definida empiricamente.

O nome *backpropagation* faz referência à retropropagação do ajuste de erros para correção dos pesos. São realizados vários ciclos de treinamento, até que o vetor de pesos seja considerado ótimo, o que significa levar ao menor erro possível da “superfície de erros” do problema. Como exemplificado na Figura 2.6, além do mínimo global também é possível encontrar mínimos locais na superfície de erros, o que pode ser um problema durante o treinamento. Além disso, em determinado ponto do treinamento a rede memoriza o processo, devido ao excesso de treinamento, referenciado na literatura como *overtraining*, causando a perda de capacidade de generalização.

O processo de aprendizagem da rede é formado por três etapas - treinamento, teste e validação - cada qual utilizando um subconjunto da série de dados.

- Treinamento: o conjunto de treinamento é a maior das parcelas, pois deve representar todas as variações ocorridas ao longo do tempo para o fenômeno em estudo.
- Teste: o conjunto de teste é apresentado ao fim de cada época de treinamento, que consiste na apresentação de todo o conjunto de treinamento, avaliando o desempenho da rede até o momento através do erro. Com sucessivos treinamentos, o erro do conjunto de teste tende a diminuir, até alcançarmos o ponto de *overtraining*, quando volta a crescer, como representado esquematicamente na Figura 2.7.
- Validação: Após realizar o treinamento, a rede é operada com o vetor de pesos ótimo, utilizando como entrada os dados do subconjunto que ainda não foi apresentado à

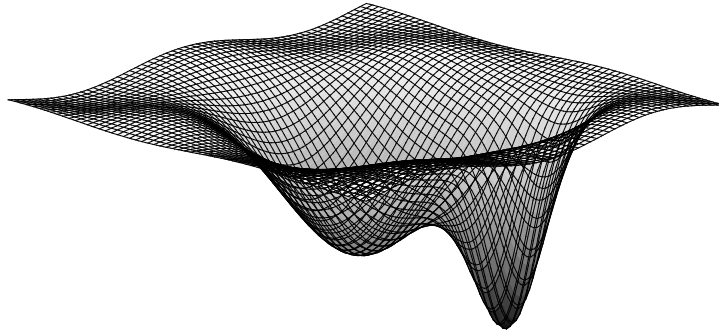


Figura 2.6: Representação de uma superfície de erros, contendo um mínimo local e um mínimo global.

rede. Os resultados obtidos pela rede são comparados então com o valor esperado, avaliando o desempenho da rede treinada.

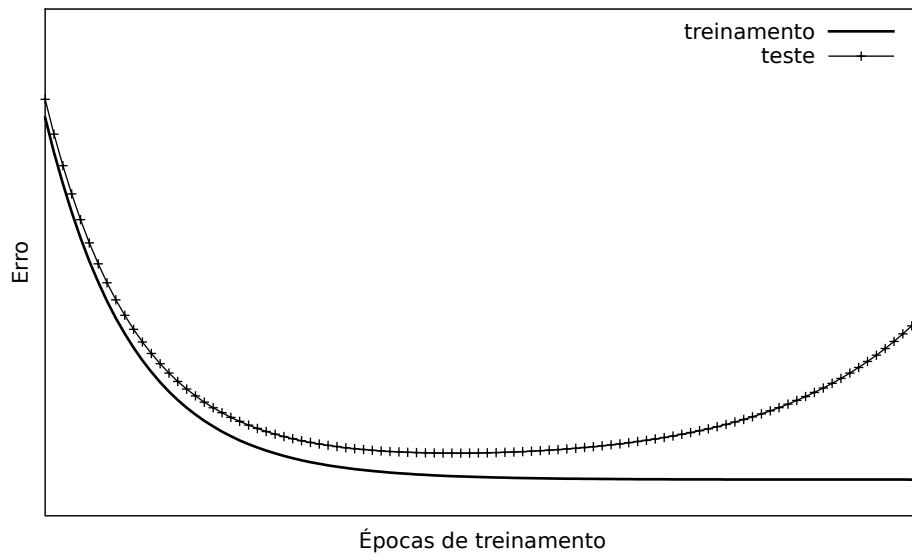


Figura 2.7: Curvas de treinamento e teste, ao longo das épocas de treinamento.

2.4 Estimativa da altura da CLA

Não há uma metodologia padrão para a determinação de h . Baseando-se em característica da CLA, pesquisadores desenvolveram diferentes definições ao longo dos anos,

e muitas vezes faz-se necessária a aplicação de métodos combinados para a obtenção de h . Contudo, esse ainda é um tema de pesquisa recorrente, posto que mesmo após anos de estudo não foi encontrada uma metodologia que apresente resultados realmente satisfatórios, considerando qualidade de resultados e também sua operacionalização.

2.4.1 Métodos de estimativa com dados de perfis de variáveis atmosféricas

Um instrumento clássico para obtenção de perfis verticais de escalares na atmosfera é a radiossonda. Como apresentado de forma mais detalhada previamente, esse é um equipamento composto por sensores acoplados a um balão meteorológico. Apesar da ampla utilização, as medições por radiossondagens possuem limitações. Os dados obtidos são pontuais, não apenas no espaço mas também no tempo.

SEIDEL *et al.* [14] apresentam um estudo comparativo com sete diferentes métodos de determinação de h a partir de radiossondagens. São eles: método da parcela, nível de máximo gradiente vertical de θ , nível de mínimo gradiente vertical de q , nível de mínimo gradiente vertical de UR , nível de mínimo gradiente vertical de refratividade (N), altura da base de uma inversão elevada de temperatura e altura do topo de uma inversão de temperatura com base na superfície. As inversões de temperatura com base na superfície foram avaliadas separadamente, dadas as características particulares de CLA's estáveis nas quais tais inversões são observadas, não sendo aplicáveis os outros seis métodos para tais casos.

Os resultados desse estudo mostram que, apesar de ocasionalmente haver correspondência entre as alturas indicadas pelos diferentes métodos (excetuados os casos de inversão de temperatura com base na superfície), há um número considerável de casos nos quais os métodos diferem em seus resultados. Em 13% dos casos não havia inversões elevadas de temperatura potencial, tampouco inversões com base na superfície. No caso dos gradientes verticais de q e UR deve ser levada em consideração a presença de nuvens, as quais podem causar incerteza nas medições do sensor. Com medições de valores de umidade bastante elevados no interior da nuvem, é comum que o método localize um gradiente acentuado no topo da nuvem, indicando equivocadamente ser este o topo da CLA. A maior correlação linear encontrada entre dois métodos foi de $r = 0.67$, entre os métodos de mínimos gradientes de umidade relativa e umidade absoluta. A correlação entre os métodos de máximo gradiente de θ e o mínimo gradiente de umidade relativa foi de $r = 0.49$.

Ao contrário das radiossondas, os equipamentos de SODAR e LIDAR tem como uma de suas vantagens a operação contínua.

O SODAR funciona com a emissão de pulsos acústicos, e suas medições se baseiam no

retroespalhamento da energia das ondas sonoras emitidas [20]. O alcance das medições varia entre diferentes equipamentos. Com altura máxima de medição de muitos equipamentos em torno de 1 km, muitas vezes as medições ficam restritas a altitudes ainda menores [14]. Já o limite inferior de medição fica a aproximadamente 50 m da superfície [1], inviabilizando seu uso para identificação de camadas muito baixas [21].

Apesar de dados de SODAR serem usados há mais de 20 anos para determinação de h , esse método ainda é visto de forma bastante controversa por muitos estudiosos da área [1].

Na aplicação de SODAR para determinação de h estudada por BEYRICH [1], são fornecidos dados de intensidade de espalhamento, vetor de velocidade horizontal, velocidade vertical (u_3) e desvio-padrão da velocidade vertical ($\sigma_{u_3}^2$). O autor ressalta que a incerteza nas medições de $\sigma_{u_3}^2$ e mesmo de u_3 chegam à mesma ordem dos próprios parâmetros, inviabilizando seu emprego direto.

BEYRICH [1] aponta também a dificuldade de estabelecer um método direto de estimativa da altura da CLA, devido às diversas forçantes atuando em diferentes graus nessa região. No caso da camada-limite estável, por exemplo, podemos ter, segundo o autor, duas situações bastante distintas, dependendo da presença de vento forte ou não. Na presença de um vento muito fraco, os gradientes são mais acentuados perto da superfície. Nesses casos, a altura estimada pelos dados de SODAR são comumente coincidentes com a altura da inversão.

Foram então definidos os diferentes critérios de determinação, de acordo com as características principais da camada a ser estudada. Para camadas bem desenvolvidas, com altura superior a 1000 metros, não é indicado o uso de SODAR pelo autor, visto que essa faixa está acima do alcance do aparelho. Tampouco foi considerado eficiente na observação de camadas muito baixas, pois seu primeiro intervalo de medição está entre 30 e 50 metros. Para situações como essa podem ser utilizados equipamentos de alta frequência, *mini-SODAR*, com primeiro intervalo de medição entre os níveis de 10 e 20 metros.

Em suas conclusões, BEYRICH [1] afirma que dados de SODAR são uma escolha confiável para estimativa de alturas da CLA entre 50 e 500 – 1000 metros, dependendo do tipo de equipamento. Os critérios para definição da altura a partir dos dados fornecidos pelo SODAR devem ser estabelecidos de acordo com a estrutura da camada e estágio de evolução. Contudo, suas medições são sensíveis a ruídos do ambiente, e a interpretação dos resultados requer um profissional com experiência [14], não sendo indicado o uso exclusivo de métodos automatizados [1].

Para estimar a altura da CLA convectiva durante a tarde, a qual usualmente ultrapassa os limites de detecção do SODAR, BEYRICH [22] emprega, com adaptações, a seguinte equação:

$$\frac{dh}{dt} = (1 + 2c_1) \frac{\langle u_3 \theta' \rangle_0}{\gamma_\theta h} + \frac{2c_2 u_*^3}{\beta \gamma_\theta h^2}. \quad (2.21)$$

Na Equação (2.21), $\langle u_3 \theta' \rangle_0$ é o fluxo de calor na superfície, β é o parâmetro de flutuabilidade g/θ , γ_θ é o *lapse rate* da temperatura potencial, e c_1 e c_2 são constantes.

BEYRICH [22] cita as incertezas nos valores das constantes, com valores propostos na literatura de $c_1 = 0.2$ e $c_2 = 2.5$, como uma das possíveis fontes de erros na aplicação da Equação (2.21), sendo proposto pelo autor o uso de dados observados de SODAR da CLA da manhã para determinação de novos valores de c_1 e c_2 melhor ajustados às observações. Os resultados obtidos pelo modelo com as constantes propostas na literatura e com as constantes ajustadas foram comparados a valores de h obtidos de radiossondagens, com emprego da definição de h como altura da base da inversão elevada de θ (z_i). Segundo o autor, os resultados confirmam a aplicabilidade do modelo adaptado, com coeficiente de correlação linear de 0.95. Contudo, cabe ressaltar que com a utilização dos valores propostos na literatura para as constantes, obteve-se um coeficiente de correlação linear médio para todos os casos de 0.94. Há ainda a limitação da proximidade de uma estação de radiossondagem, a fim de obter dados de um perfil de temperatura inicial.

Assim como o SODAR, outra ferramenta de sensoriamento remoto, o LIDAR, vem se destacando na estimativa da altura da CLA. Suas características e os critérios para determinação de h , contudo, são diferentes. Inversões de temperatura na atmosfera são caracterizados por gradientes acentuados de escalares, e é na observação de gradientes de vapor d'água e aerossóis que se baseia o método de determinação de h do LIDAR.

Utilizando um equipamento com detecção de vapor d'água, HENNEMUTH e LAMMERT [2] apresentam uma comparação entre seus resultados e métodos clássicos de determinação de h a partir de radiossondagens. O foco do trabalho é a determinação da camada limite convectiva e também a observação de camadas residuais. Isso se deve às limitações de detecção do equipamento, que inviabiliza sua utilização no estudo de camadas limite muito baixas, como usualmente se observa em camadas limite noturnas [2].

A determinação de h pelo LIDAR é feita aplicando-se duas metodologias diferentes. Na primeira foi calculada a variância do sinal ao longo de um período de 30 min, e a segunda resultou em valores instantâneos de h a cada 10 s, a partir do mínimo gradiente do perfil de espalhamento. A altura resultante desse método foi considerada como uma altura média \bar{h} para o período em questão. Segundo os autores, o método de gradiente é bastante sensível a ruídos, e antes de sua aplicação, foi definido um limite vertical de 350 m em torno do valor \bar{h} já obtido para o período. A mesma metodologia foi utilizada por LAMMERT e BÖSENBERG [23], num estudo comparativo entre a utilização dos métodos de gradiente e variância, com os mesmos dados apresentados por HENNEMUTH e LAMMERT [2]. As diferenças relativas são $\lesssim 100$ m, correspondendo a $\lesssim 5\%$, e os métodos

foram considerados equivalentes pelos autores de ambos os artigos. A comparação entre os métodos foi viabilizada considerando-se uma média das alturas instantâneas obtidas ao longo do período de 30 min.

Em quatro campanhas para obtenção de dados, foram lançadas de duas a quatro radiossondas por dia, a distâncias de 4 km a 50 km dos pontos de operação do LIDAR. A determinação de h a partir de radiossondagens foi automatizada, utilizando os métodos da parcela, de máximo gradiente de θ e mínimos gradientes de UR e umidade absoluta (UA), com resolução entre 25 m e 50 m. Segundo os autores, os métodos foram concordantes em muitos casos, mas houve situações com múltiplos gradientes nas quais foi observada grande divergência nos valores de h estimados.

Ao comparar os valores de h obtidos pelo LIDAR, e os valores obtidos pelo método da parcela, foi encontrado um desvio médio de 400 m, com valores obtidos pelo LIDAR usualmente maiores. Já quando comparado ao método de mínimo gradiente de umidade, houve uma boa concordância entre ambos, com número considerável de casos de subestimação, se tomarmos os valores obtidos a partir da radiossondagem como referência. Contudo, o método do mínimo gradiente de umidade determinou muitas vezes a altura da camada residual, principalmente no período da manhã. Em dias de céu limpo o desempenho do LIDAR foi bastante satisfatório segundo os autores, com desvio médio em relação aos valores obtidos a partir de radiossondagens de ± 200 m.

Em [24], ZELAYA-ÁNGEL *et al.*, assim como nos casos de aplicação de LIDAR e SODAR, os autores buscam a característica do topo da CLA de retenção de substâncias, mas nesse caso é feita a análise de concentração de eteno, em lugar de vapor d'água e material particulado. Os autores aplicaram um modelo fotoquímico euleriano baseado nas equações de conservação de massa, aplicadas a um volume finito em forma de “caixa”. A base da caixa tem área $\Delta x \Delta y$, e a altura da caixa é a própria altura da CLA. As medições de eteno atmosférico na Cidade do México são aplicadas ao modelo euleriano para determinar os coeficientes da função de Boltzmann, utilizada então para determinação da evolução da altura da CLA.

2.4.2 Métodos de estimativa a partir de dados de superfície

Com uma abordagem diferente, LI [13] apresenta uma parametrização da altura da CLA estável com emprego de redes neurais artificiais. Os dados utilizados nesse estudo são a saída de 68 simulações de um modelo LES (*large eddy simulation*). A limitação vertical de interesse foi definida em 500 metros, e o critério de definição do topo da CLA estável escolhido foi o nível do mais baixo máximo na velocidade do vento da própria simulação do modelo LES. Tal critério implicou em uma limitação no horário a ser estudado, definido entre 00 e 06 h, pois observações dos perfis de vento do LES mostram que o jato de baixo

nível começa a se desenvolver após as 00 h.

O conjunto de dados de entrada foi então separado em treinamento (60%) e validação (40%), e o objetivo foi definir uma formulação linear para altura da CLA estável. Como entrada foram utilizados a velocidade de atrito u_* , fluxo de fluatibilidade B_s , frequência de Brunt-Vaisala N , o parâmetro de Coriolis f e a altura de rugosidade z_0 , combinados de diferentes maneiras, a fim de definir uma configuração ótima de dados de entrada para parametrização. A combinação de dados de entrada com melhores resultados utilizou todos os parâmetros acima mencionados, e uma avaliação detalhada dos resultados das diversas combinações mostrou que o parâmetro mais importante para determinação da altura da CLA estável para esse estudo é u_* .

O desempenho da rede neural foi também comparado ao emprego das seguintes parametrizações:

$$h = C_n \frac{u_*}{f}, \quad (2.22)$$

$$h = C_{sr} \frac{u_*^2}{-f B_s}^{1/2}, \quad (2.23)$$

$$h = C_{ir} \frac{u_*}{f N}^{1/2}, \quad (2.24)$$

$$h = -C_s \frac{u_*^3}{B_s}, \quad (2.25)$$

$$h = C_i \frac{u_*}{N}, \quad (2.26)$$

Essas parametrizações, existentes na literatura, representam relações entre os parâmetros utilizados como dado de entrada pela rede e a altura da camada limite estável [13], e com C_n , C_{sr} , C_{ir} , C_s e C_i sendo coeficientes empíricos.

Dentre as parametrizações avaliadas, a que apresentou melhores resultados foi a expressa pela Equação (2.22). A inserção apenas de u_* e f na rede neural resultou em um valor de raiz do erro médio quadrático de 15.59 m, contra 22.15 m encontrado com o emprego da Equação (2.22), que utiliza os mesmos parâmetros na determinação de h .

PELLICIONI *et al.* [25] empregaram uma rede neural artificial para estimar a altura da camada de mistura e o valor do comprimento de Monin-Obukhov, utilizando informações de superfície oriundas de uma estação meteorológica.

A base de dados para a operação da rede foi obtida através de equações baseadas em parametrizações, aplicando-se os próprios valores de superfície medidos na estação. Na estimativa de L , foram empregadas as equações de Monin e Obukhov para L (2.7),

θ_{v*} (2.8) e u_* (2.9), e para as funções de estabilidade (2.10, 2.11 e 2.12), resolvidas de maneira iterativa até convergência. Em casos de não convergência, L foi estimado através de uma formulação de Panofsky [25]. A altura da camada de mistura foi obtida através da parametrização proposta por PANOFSKY [17],

$$h = \frac{T_h - T_{sunrise}}{\Gamma_d - \Lambda}, \quad (2.27)$$

para camadas instáveis. Para condições estáveis ou neutras foi empregada a equação

$$h = \frac{u_*}{|f|} \left\{ \frac{1}{A_0} + \left[\left(\frac{u_*}{|f|k} \right)^{1/2} \frac{1}{C_H} \left(\frac{1}{L} \right)^{1/2} \right]^{-1} \right\}, \quad (2.28)$$

a qual é uma parametrização proposta por ZILITINKEVICH [15], e na qual A_0 e C_H são constantes adimensionais.

De posse de uma base de dados, foi empregada uma rede neural artificial do tipo perceptron multicamadas, com treinamento pelo método de *backpropagation* com regra de aprendizado supervisionado. Os parâmetros considerados significantes para entrada na rede foram: hora do dia (*hora*), temperatura (T_{hora}), temperatura uma hora antes (T_{hora-1}), velocidade horizontal do vento (U), desvio padrão do vento (σ_U) e radiação solar (SR).

Os parâmetros a serem estimados foram h e $1/L$, e a melhor configuração da rede obteve valores de correlação de $r = 0.81$ e $r = 0.82$, para h e $1/L$ respectivamente.

Em outro estudo, FREIRE [26] faz uma avaliação de algumas teorias de CLA, usando para tal dois conjuntos de dados, de dois sítios distintos: um obtido por radiossondas no Kansas, e outro pelo uso de um veículo aéreo não-tripulado (VANT) em Tijucas do Sul. Como parte desse trabalho foi observado o modelo de crescimento da CLA salto de ordem zero, o qual utiliza como entrada dados de superfície (fluxo de calor e velocidade de atrito do vento) e o gradiente vertical de temperatura potencial acima da inversão. Com esse modelo foi estimada a altura da CLA diurna em três dias de medição em Tijucas do Sul, seguindo a seguinte equação:

$$\frac{dz_i}{dt} - u_{3s} = \frac{\overline{u'_3 \theta'_v} |_0 [(1 + 2A)z_i - 2B\kappa L]}{\gamma_\theta z_i^2} \quad (2.29)$$

Os valores de erro obtidos são apresentados na Tabela (2.1).

Para o conjunto de dados obtido por radiossonda no Kansas são apresentados apenas os valores médios dos erros, considerando todo o conjunto. Esses resultados são apresentados na Tabela (2.2). Para esse conjunto de dados foi considerada a presença de camada residual, sendo então usados dois γ_θ 's distintos, um para a região de camada residual e

Tabela 2.1: Erros obtidos pelo modelo de crescimento em Tijucas do Sul.

	Dia 1	Dia 2	Dia 3
MAE	105.67	49.40	25.16
RMSE	196,41	89,73	50,37
r	0.98	0.94	0.81

um para a atmosfera livre.

Tabela 2.2: Erros obtidos pelo modelo de crescimento no Kansas.

	Modelo com 1 γ_θ	Modelo com 2 γ_θ 's
MAE	761.06	195.33
MAB	-753.92	130.34
RMSE	947.68	256.85

Na revisão de literatura não foram encontrados métodos para estimativa de h com base apenas em dados de observação de superfície. Outras metodologias apresentadas nesta seção para estimativa de h são avaliadas de maneira subjetiva por seus autores e, de maneira geral, a correlação linear é o único parâmetro objetivo de avaliação apresentado.

Capítulo 3

Materiais e Métodos

3.1 Base de dados

O treinamento e a operação de uma RNA exige uma base de dados confiável, com número suficiente de pares entrada-saída. Esses pares são os vetores de entrada, que reúnem os parâmetros considerados necessários e relevantes para a resolução do problema, e saída, com as variáveis que se quer estimar. Nosso vetor saída tem apenas uma componente, a altura da CLA. Já o vetor de entrada será variável, com o teste de diferentes combinações dos parâmetros escolhidos - para cada configuração, um diferente vetor de entrada, com nomeação específica para sua identificação.

Uma base de dados dentro das características citadas e com valores de h estimados não foi encontrada na literatura, sendo então necessária a obtenção dessa série especificamente para esse estudo. Para tanto, optamos pela determinação de h a partir de perfis medidos por radiossondagens.

3.1.1 Dados de radiossondagens

Dentre vários projetos com dados de radiossondagens disponíveis, escolhemos utilizar dados do *National Oceanic and Atmospheric Administration* (NOAA), disponibilizados gratuitamente pelo *SPARC* (*Stratosphere-troposphere Processes and their Role in Climate*) em seu endereço eletrônico (<http://www.sparc-climate.org/>) em formato de arquivo de texto. Duas características foram fundamentais na escolha do projeto: a alta resolução das medições e o grande número de estações espalhadas por todo o território dos Estados Unidos da América, possibilitando a escolha de estações com diferentes características geográficas.

Ao todo são disponibilizados dados de 90 estações, no período de 1998 a 2011, com frequência de medições dos sensores de 6 s. A partir de 2005 teve início uma transição para

uso de sensores com frequência de medição de 1 s. As estações escolhidas estão entre as que passaram pela transição até o ano de 2011, pois a frequência de medição de 1 s possibilita uma representação mais detalhada dos perfis. Considerando uma velocidade média de ascensão de 4 m/s, a sondagem de frequência de 6 s faz medições a aproximadamente cada 24 m, enquanto a com frequência de 1 s faz medições a cada 4 m.

Usualmente são lançadas duas radiossondas por dia, nos horários de 00 e 12 UTC. Esporadicamente há dias com lançamentos disponíveis a cada 3 horas. Identificados por estação, ano, mês, dia e hora do lançamento, há arquivos distintos contendo dados brutos e já processados pelo *SPARC* para correção de temperatura devido à radiação solar, e também correções de falhas na obtenção de dados de velocidade. Mesmo com as correções realizadas pelo *SPARC* foi necessário um segundo processamento e controle de qualidade, pois há radiossondagens com falhas na medição de umidade em diversos níveis, sem que haja qualquer indicativo com valores de referência (-99.9) como em outros casos. Há também falhas na medição da altura da radiossondagem pelo equipamento de GPS, também não indicadas nos arquivos ou na documentação do projeto. Sondagens com falhas de velocidade nos primeiros 50 m ou falhas de umidade nos primeiros 4 km foram desconsideradas, procedimento justificado na Seção 3.1.2.

As variáveis consideradas relevantes para esse estudo retiradas desses arquivos processados são listadas na tabela 3.1.

Tabela 3.1: Dados retirados das radiossondagens de alta resolução.

Dado	Unidade
Temperatura	°C
Umidade Relativa	%
Pressão	hPa
Componentes horizontais de velocidade do vento	m/s
Altura Geopotencial	gpm

3.1.2 Procedimentos para análise

Como em todo trabalho com dados de campo, estamos sujeitos às falhas que podem ocorrer nos sensores, fazendo-se necessário um pré-processamento para controle de qualidade.

Foram utilizados para o desenvolvimento do projeto dados de níveis inferiores a 4 km de altura, considerando os valores geralmente encontrados para CLA na literatura. Radiossondagens sem medição de velocidade foram descartadas, mesmo sendo possível a estimativa de h a partir dos perfis de umidade e temperatura, pois os dados de velocidade do primeiro nível serão inseridos posteriormente como entrada da rede neural, e as velocidades de até 50 m são utilizadas na estimativa da altura de rugosidade. Outro problema encontrado foi a repetição da altura da radiossonda fornecida pelo GPS, em níveis consecutivos. Tais níveis e os dados referentes a eles foram excluídos, mantendo-se apenas o primeiro registro com a determinada altura em repetição.

Foram utilizados dados de temperatura (T), umidade relativa (UR), componentes horizontais do vento (U_1 e U_2), pressão (p) e altura geopotencial (h_{gpt}), para cada nível, para calcular outras variáveis de interesse, sendo elas: temperatura potencial virtual em K (θ_v) e suas derivadas de primeira e segunda ordens em relação à altura, umidade específica em g/g (q) e suas derivadas de primeira e segunda ordens em relação à altura, módulo da velocidade horizontal do vento (U_H) em m/s, e número de Richardson volumétrico (Ri_b).

Primeiramente, foi utilizada uma média móvel de 3 pontos para obter perfis suavizados de θ_v e q , e também suas respectivas derivadas de primeira e segunda ordens. Os perfis suavizados e suas derivadas foram utilizados juntamente aos perfis de dados brutos, a fim de auxiliar na determinação da altura da CLA. Foi calculado o número de Richardson volumétrico para as sondagens, utilizando tanto os dados brutos (Ri_b), quanto os dados suavizados (Ri_{bs}).

A partir dos dados de U_H calculados, utilizamos o método dos mínimos quadrados para um ajuste de reta, com $y = \ln(h_{gpt})$ e $x = U_H$, com intuito de estimar uma altura de rugosidade de referência z_{0r} . Para o ajuste de reta foram consideradas apenas perfis de velocidade monotonicamente crescentes com a altura, num nível máximo de 50 m e um número mínimo de 6 pontos. O valor médio de z_{0r} foi utilizado como valor inicial numa primeira estimativa dos parâmetros u_* , θ_{v*} e L , aplicando as Equações (2.7) a (2.12), num algoritmo iterativo. O critério de parada utilizado foi um erro de $er \leq 1 \times 10^{-3}$ entre os valores de L das últimas duas iterações. O valor de L estimado foi aplicado como critério de estabilidade, e casos com $-0.1 < z/L < 0.1$ foram considerados próximos à neutralidade, condição necessária para a aplicação da Equação 2.4. O valor de z utilizado foi a altura da camada superficial, estimada como $h/10$, o que limitou o número de registros. Os valores de z_{0r} referentes a esses casos foram utilizados para uma nova média, e o processo foi repetido até que todos os casos considerados para a última média de z_{0r} atendessem ao critério de estabilidade adotado, e pudessem ser utilizados para a determinação de um valor de z_0 para caracterização dos sítios.

3.1.3 Caracterização dos sítios analisados

As estações foram escolhidas principalmente por sua localização geográfica, visando representar as diferenças climáticas presentes em regiões distintas dos Estados Unidos, e também evitar estações litorâneas ou muito próximas (menos de 100 km) de grandes massas de água, visto que as teorias aplicadas para CLA consideram homogeneidade horizontal, e essas regiões têm comportamento bastante particular. Com base nesses critérios, foram selecionadas três estações: Reno, no estado de Nevada, Amarillo, no Texas, e Peachtree, na Georgia.

Amarillo

Em Amarillo, a estação meteorológica está localizada no aeroporto internacional *Rick Husband*, na periferia da cidade. As informações de precipitação média e também as médias mensais de temperaturas máximas, médias e mínimas, ao longo de 30 anos são apresentadas na Figura 3.2. Com uma precipitação anual média de 517 mm, se encontra em uma região de vegetação rasteira. A Figura 3.1 traz uma imagem de satélite da localização da estação, e também do entorno. A população estimada no ano de 2013 era de aproximadamente 196 mil habitantes, e suas construções são de pequeno e médio porte.

Peachtree

Peachtree é uma cidade pequena, com 35000 habitantes. Fundada em 1959, a cidade é considerada um exemplo de urbanismo.

A estação de Peachtree fica localizada no aeroporto da cidade, *Peachtree City - Falcon Field*, em área urbana de pouca ocupação, com pequenas construções e bastantes áreas com cobertura vegetal densa. No entorno do aeroporto há uma área de reserva para preservação ambiental, com trilhas para recreação e árvores de grande porte, como é possível observar na Figura 3.3.

Na Figura 3.4 são apresentadas as médias mensais de precipitação, e temperaturas mínima, média e máxima, utilizando dados do período de 1981 a 2010. Peachtree é uma cidade bastante úmida, com precipitação anual média de 1270 mm.

Reno

Localizada na região oeste, a cidade de Reno é a mais seca dentre as três localidades escolhidas para esse trabalho, como mostra a Figura 3.5, com médias de precipitação mensal, e médias de temperaturas máxima, média e mínima medidas na cidade. Sua precipitação média é de apenas 210 mm por ano, aproximadamente 6 vezes menor que a precipitação anual média para a cidade de Peachtree.

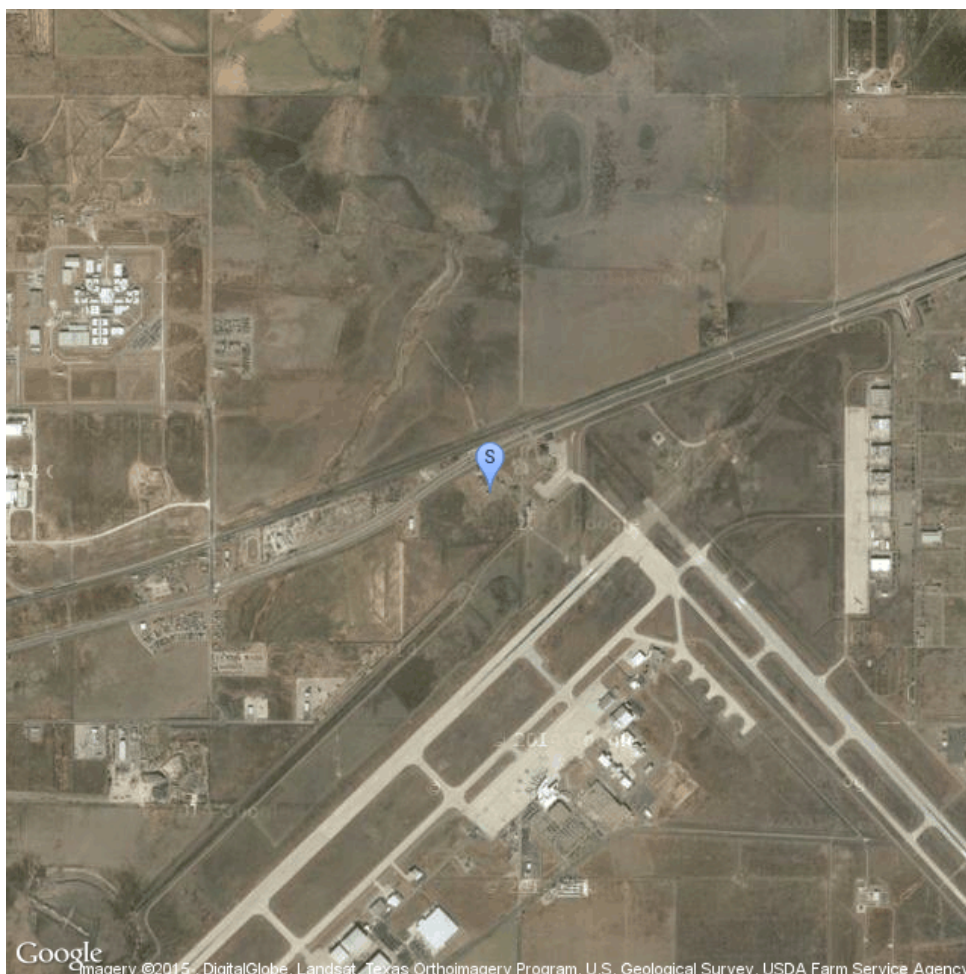


Figura 3.1: Imagem de satélite da localização da estação de Amarillo.
Fonte: API do Google Maps.

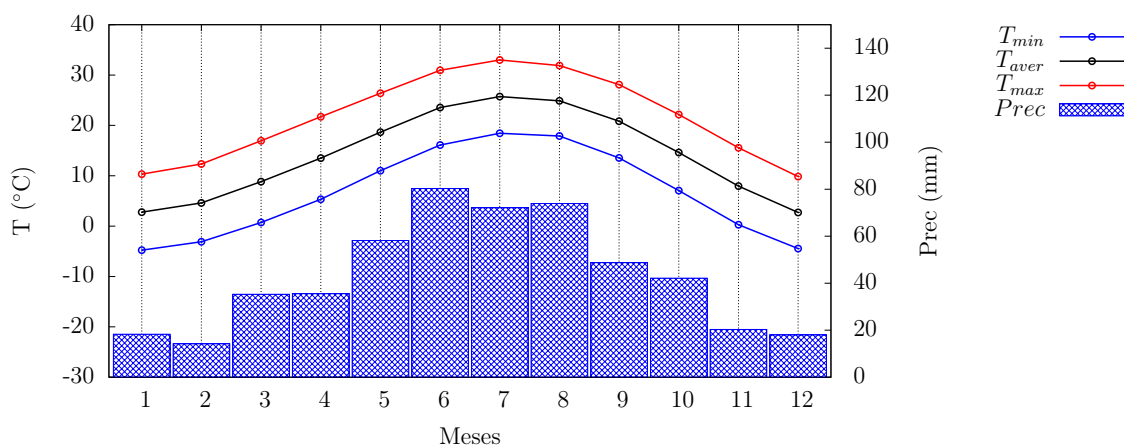


Figura 3.2: Médias ao longo de 30 anos de precipitação e temperatura para Amarillo.

A estação de Reno se encontra na Universidade Comunitária de Truckee Meadows, também na periferia da cidade, apresentada na Figura 3.6. Com urbanização predomi-

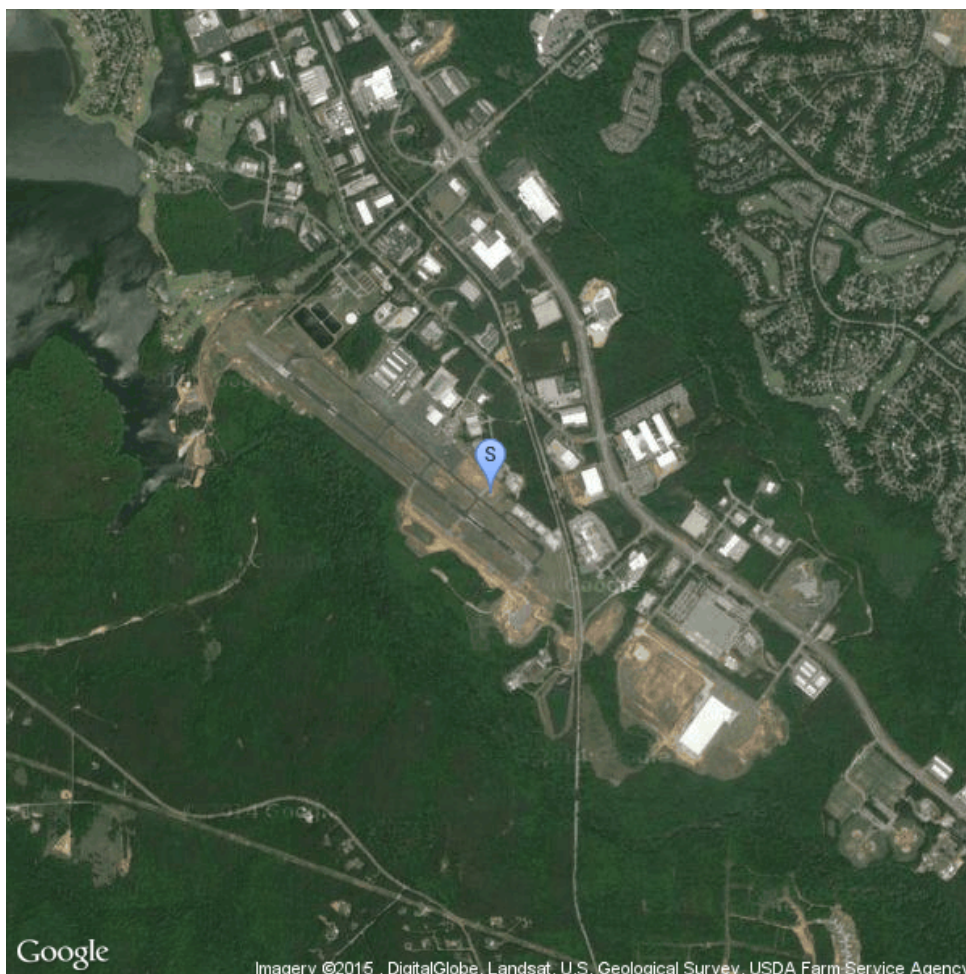


Figura 3.3: Imagem de satélite da localização da estação de Peachtree.
Fonte: API do Google Maps.

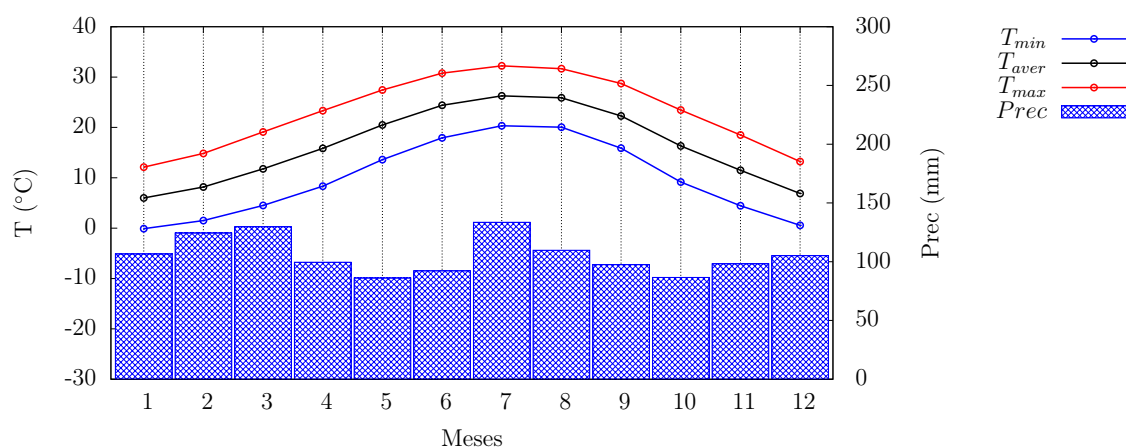


Figura 3.4: Médias ao longo de 30 anos de precipitação e temperatura para Peachtree.

nante de casas e construções de pequeno porte, Reno está situada em uma área desértica, com vegetação rasteira escassa e poucas árvores com origem na ação antrópica, com

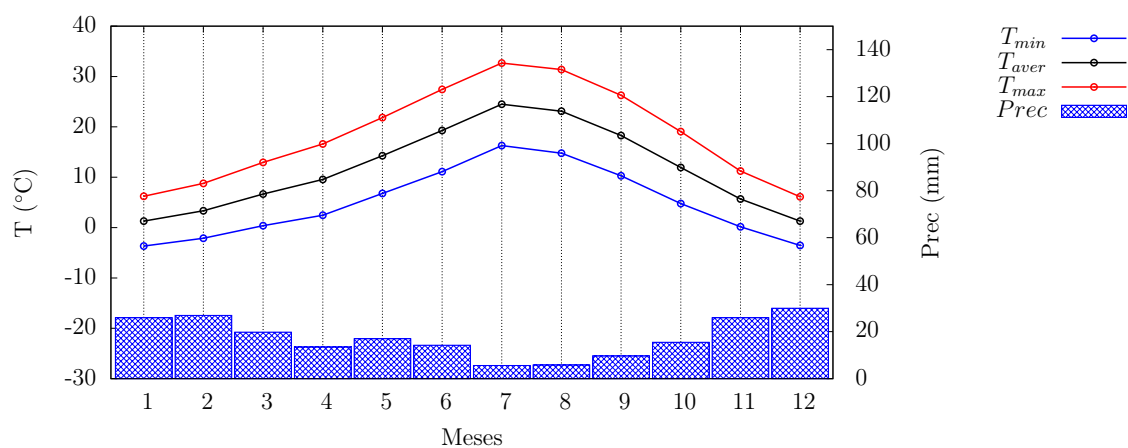


Figura 3.5: Médias ao longo de 30 anos de precipitação e temperatura para Reno.

caráter paisagístico e de urbanização.

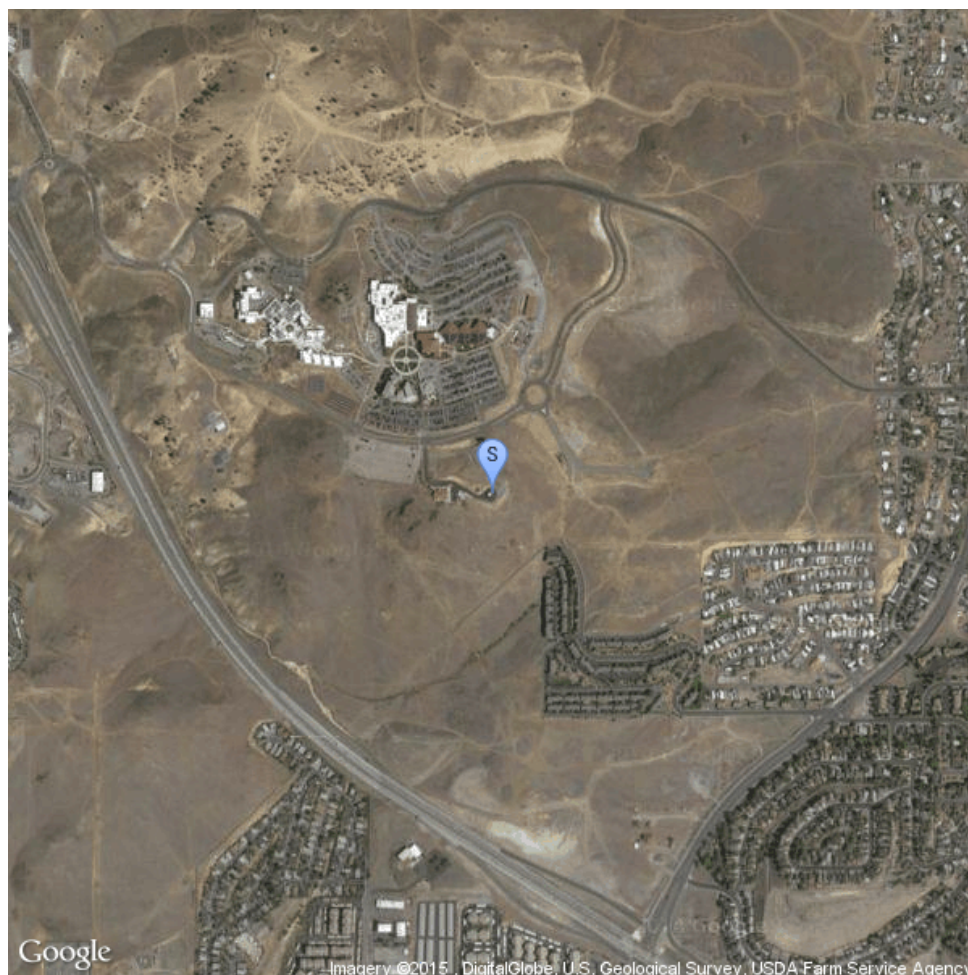


Figura 3.6: Imagem de satélite da localização da estação de Reno.
Fonte: API do Google Maps.

3.1.4 Estimativa da altura da CLA a partir de radiossondagens

As estimativas de h foram realizadas por observação dos gráficos e dados, seguindo uma série de critérios. Dos métodos apresentados por SEIDEL *et al* [14], usamos o máximo gradiente vertical de θ_v e o mínimo gradiente de q de forma automatizada, e também o nível do primeiro máximo de velocidade do vento para inversões baseadas na superfície.

Para cada sondagem, critérios diferentes resultaram em alturas diferentes, e a escolha por um ou outro critério, ou mesmo por uma outra altura não indicada por nenhum deles, foi subjetiva. Quando ambos os métodos para inversões elevadas (dq/dz e $d\theta_v/dz$) indicaram alturas próximas corretamente, o critério foi conservador, optando pela menor altura, visto que o intuito do uso dos resultados é a aplicação em modelos de dispersão atmosférica de poluentes.

3.2 Parametrização da altura da CLA

Há várias equações para parametrização da altura da CLA disponíveis na literatura, como previamente mencionado, mas apenas para casos nos quais a atmosfera é considerada estável. Em [25] é utilizada uma parametrização para altura da camada de mistura, em situação instável, proposta por Panofsky, Equação (2.27), reapresentada abaixo:

$$h = \frac{T_h - T_{sunrise}}{\Gamma_d - \Lambda}.$$

Os dados usados por PELLICIONI [25] para emprego de uma rede neural artificial para estimativa da altura da camada de mistura foram obtidos de uma estação meteorológica, sendo todos referentes ao nível da superfície. Com intuito de observar a adequação da Equação (2.27) para estimativa da altura da camada de mistura das sondagens utilizadas nesse trabalho, propusemos a aplicação da mesma parametrização utilizando os dados de temperatura e umidade obtidos dos perfis.

A parametrização de Panofsky supõe o conhecimento do *lapse rate* da camada limite como um todo, tanto no momento do nascer do sol quanto na hora em que se pretende fazer a estimativa de h . O uso apenas de dados próximos à superfície não é representativo da variação média de temperatura na camada em toda a sua extensão, e no trabalho de PELLICIONI usado como referência não há informação sobre quais níveis foram utilizados para suas estimativas, tampouco sobre a substituição do *lapse rate* observado pelo adiabático seco ou úmido, havendo apenas a nota de que não são disponíveis dados de altitude, e sim apenas dados de superfície. Isto posto, foram propostas duas diferentes configurações para comparação. Na primeira, foram usados os dados de dois níveis da radiossondagem próximos à superfície, o primeiro nível de medição da radiossonda, e nível

mais próximo da altura de referência $z = 10$ m. Para a segunda configuração, manteve-se o primeiro nível de medição, mas o segundo nível utilizado foi o referente à altura da CLA, sendo utilizado o valor de h estimado a partir dos perfis das radiossondagens.

Para aplicação dessa parametrização, são requeridos dados do momento do nascer do sol, e as sondagens disponíveis são para os horários de 12 e 00 UTC, e os lançamentos são realizados cerca de meia hora antes do horário UTC indicado, o que equivale aos horários locais de 03 : 30 h, 05 : 30 h e 06 : 30 h para Reno, Amarillo e Peachtree, respectivamente. Considerando tal fato, justificamos o emprego da parametrização apenas para a estação de Peachtree, em alguns dias do mês de junho, quando a hora de lançamento (06 : 30 h) é o mais próxima possível do horário do nascer do sol na cidade, informado como 06 : 28 h no horário local.

3.3 Treinamento e validação da Rede Neural Artificial

O programa para treinamento da rede neural artificial empregada foi escrito pelo Prof. Ricardo Carvalho de Almeida, e é composto por dois códigos escritos em linguagem de programação FORTRAN, um para preparação dos conjuntos de treinamento, teste e validação, e outro para execução de treinamento e operação.

A preparação dos dados consiste em sua normalização dentro do intervalo $[-1, +1]$, a fim de evitar que parâmetros de ordem de grandeza diferentes prejudiquem o aprendizado da rede, e a separação dos conjuntos de treinamento, teste e validação. Os dados de entrada serão multiplicados pelos pesos sinápticos e somados em cada neurônio da camada intermediária. Valores de ordem de grandeza maior tendem a mascarar a importância de valores de ordem de grandeza muito menor, como os dados de pressão comparados aos valores de rugosidade, por exemplo. Daí a importância da normalização dos dados de entrada. Os conjuntos de treinamento, teste e validação são compostos seguindo respectivamente as porcentagens de 75%, 15% e 10% do total de pares entrada-saída, de maneira aleatória.

Os conjuntos previamente separados são utilizados pelo segundo código para treinamento e validação da rede. Esse programa consiste numa rede do tipo *multilayer perceptron*, com uma camada intermediária. O treinamento da rede é realizado pelo método de *backpropagation* por batelada. Um número máximo de iterações é definido subjetivamente, e um gráfico comparativo entre o erro médio de cada iteração para os conjuntos de treinamento e teste é utilizado como referência para essa definição. Mesmo em casos de *overtraining*, o programa salva a matriz de sinapses considerada ótima e a imprime em um arquivo texto para uso posterior na validação da rede. Como dados de saída temos

também os índices estatísticos empregados para avaliação de qualidade dos resultados.

Para uma mesma matriz de sinapses podemos obter resultados consideravelmente diferentes se utilizados dois conjuntos diferentes de validação. Dessa forma, é recomendado o emprego de uma validação cruzada, com 10 rotinas de preparação de dados, treinamento e validação diferentes para cada configuração testada, sendo então calculada a média dos índices estatísticos de cada uma das 10 validações.

Cada configuração é composta por diferentes combinações de parâmetros no vetor de entrada. Como dados escolhidos para compor os diferentes vetores de entrada estão parâmetros meteorológicos e informações para caracterização dos sítios, além de informações sobre a última medição de h disponível. A caracterização dos sítios é feita com o emprego de z_0 , e também com a informação de sua localização geográfica, expressa pela latitude (lat). Como parâmetros meteorológicos de superfície foram empregados dados do primeiro nível de medição da radiossonda como representação de dados que poderiam ser obtidos em uma estação meteorológica comum, sendo eles temperatura (θ_0), pressão (P_0), umidade relativa (UR_0) e velocidade horizontal do vento (U_{H0}), todos no nível da superfície, além de u_* . Muitas estações meteorológicas não possuem medição de dois níveis de temperatura para obtenção de u_* pela teoria de Monin-Obukhov. Contudo a inclusão de medição de temperatura em um segundo nível não seria difícil, e o parâmetro u_* é de extrema importância nos estudos de CLA. Dados de altitude foram utilizados com intuito de representar fenômenos de escala sinótica, e foram retirados da última radiossondagem anterior, a qual estaria disponível numa situação de operação real. Tais parâmetros são as componentes horizontais do vento para os níveis de 850 hpa (U_{1-850} e U_{2-850}) e 500 hpa (U_{1-500} e U_{2-500}), e suas respectivas alturas geopotenciais (h_{g850} e h_{g500}), assim como a umidade relativa para o nível de 700 hpa (UR_{700}). A última altura h_{old} medida, num instante $t - \Delta t$ antes da altura que se quer estimar h , medida no instante t , e o tempo Δt decorrido entre as duas medições também foram incluídos como parâmetros de entrada. As últimas informações consideradas indicam o período do dia *hora*, e a estação do ano *est*. O parâmetro *hora* não indica a hora exata de lançamento da radiossonda, mas sim se a hora se encontra no período de 23 a 11 h, representado pelo valor -1 , ou no período de 11 a 23 h, representado pelo valor $+1$. A rede não é capaz de interpretar o padrão cíclico do sistema horário, no qual 23 h é seguido por 00 h, e a separação em períodos representados no sistema binário evita esse problema. O mesmo é feito com as estações do ano. Para representá-las foram utilizados 4 neurônios, como apresentado na Tabela 3.2.

Com a base de dados já definida, demos início ao treinamento e validação da RNA. As configurações testadas para treinamento e validação da RNA foram divididas em três blocos de experimentos, de acordo com as características dos parâmetros de entrada, sendo eles:

Tabela 3.2: Identificação das estações do ano para entrada na RNA

Meses	N1	N2	N3	N4
Dez-Fev	1	-1	-1	-1
Mar-Mai	-1	1	-1	-1
Jun-Ago	-1	-1	1	-1
Set-Nov	-1	-1	-1	1

- *st*: parâmetros de estações meteorológicas de superfície, hora, mês e caracterização do sítio;
- *sth*: parâmetros de estações meteorológicas de superfície, hora, mês, caracterização do sítio e o valor da última altura da CLA medida antes do momento em que se deseja estimar h , bem como o intervalo temporal decorrido;
- *sta*: parâmetros de estações meteorológicas de superfície, hora, mês, caracterização do sítio, parâmetros de altitude, valor da última altura da CLA medida antes do momento em que se deseja estimar h e o intervalo temporal decorrido.

Os sufixos *-d* e *-n* representam, respectivamente, dia e noite, referindo-se aos testes nos quais o conjunto de dados foi inserido separadamente, considerando a caracterização realizada previamente para os dois períodos do dia, visando avaliar o desempenho de duas redes especializadas.

A Tabela 3.3 apresenta todas as configurações testadas, com os respectivos parâmetros que as constituem.

Algumas configurações que apresentaram melhor desempenho tomando-se como referência índices estatísticos apresentados na Seção 3.4 foram submetidas a uma modificação, com os valores de h sendo substituídos por $\ln(h)$. Essa modificação é comumente utilizada na literatura quando os dados apresentam características de distribuição muito assimétricas, como é o caso dos valores de h , e pode ser observado na Figura 3.7. A aplicação do logarítmo tende a aproximar a série de dados a uma distribuição normal, e o gráfico da distribuição de $\ln(h)$ é representado na Figura 3.8.

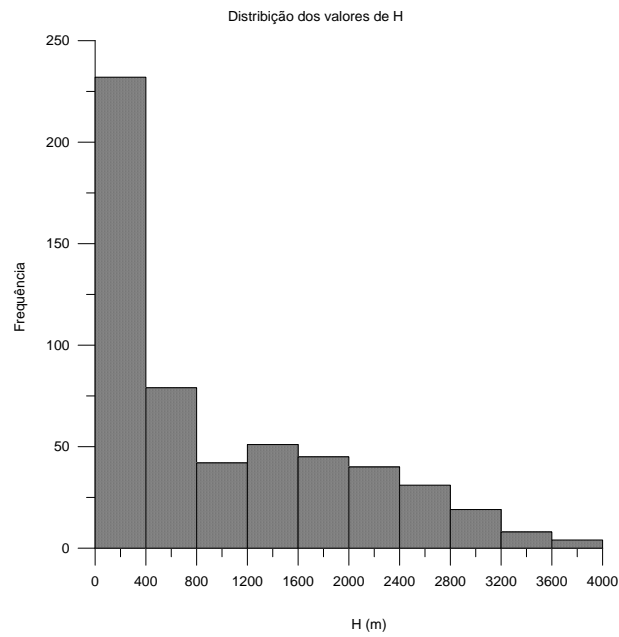


Figura 3.7: Distribuição dos valores de h .

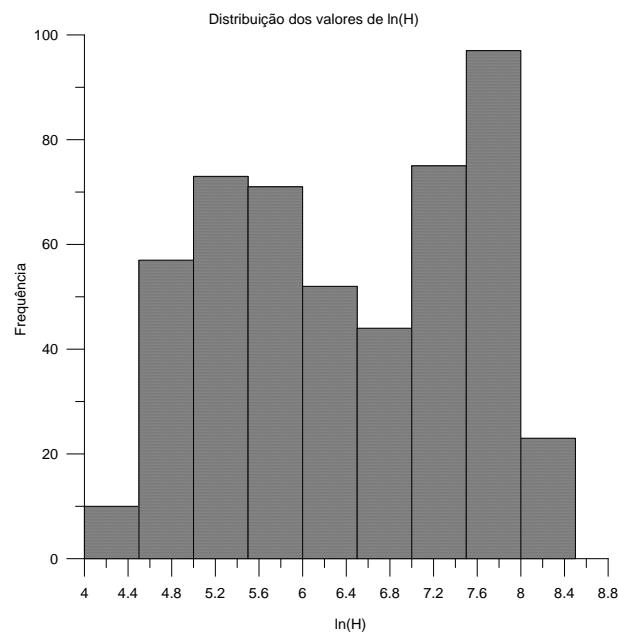


Figura 3.8: Distribuição dos valores de $\ln(h)$.

Tabela 3.3: Diferentes configurações de entrada para treinamento e validação da RNA.

	<i>hora</i>	<i>est</i>	θ_0	P_0	z_0	UR_0	U_*	U_{H0}	<i>lat</i>	Δt	<i>h_{old}</i>	U_{1-850}	U_{2-850}	h_{g850}	U_{1-500}	U_{2-500}	h_{g500}	UR_{700}
st1	✓	✓	✓	✓	✓	✓	✓	✓	✓									
st2	✓	✓	✓	✓	✓	✓	✓	✓										
st3	✓	✓	✓	✓	✓	✓	✓		✓									
st4	✓	✓	✓	✓	✓	✓		✓	✓									
st5	✓	✓	✓	✓	✓		✓	✓	✓									
st6	✓	✓	✓	✓		✓	✓	✓	✓									
st7	✓	✓	✓		✓	✓	✓	✓	✓									
st8	✓	✓		✓	✓	✓	✓	✓	✓									
st9	✓	✓				✓	✓	✓	✓									
st10	✓	✓	✓			✓	✓	✓	✓									
st11			✓	✓	✓	✓	✓	✓										
st12		✓	✓	✓	✓	✓	✓	✓										
st13	✓		✓	✓	✓	✓	✓	✓										
st14			✓	✓	✓	✓	✓	✓	✓									
st15	✓	✓	✓	✓	✓	✓	✓	✓										
st16	✓	✓	✓	✓	✓	✓	✓	✓	✓									
st1-n/st1-d		✓	✓	✓	✓	✓	✓	✓	✓									
st2-n/st2-d		✓	✓	✓	✓	✓	✓	✓										
st3-n/st3-d		✓	✓	✓	✓	✓	✓		✓									
st4-n/st4-d		✓	✓	✓	✓	✓		✓	✓									
st5-n/st5-d			✓	✓	✓	✓	✓	✓	✓									
st6-n/st6-d			✓	✓	✓	✓		✓	✓									
st7-n/st7-d		✓	✓		✓	✓	✓	✓	✓									
sth1											✓							
sth2									✓									
sth3	✓	✓			✓	✓	✓											
sth4	✓	✓			✓	✓	✓			✓								
sth5	✓	✓			✓	✓	✓		✓	✓								

Tabela 3.3: Diferentes configurações de entrada para treinamento e validação da RNA.

	<i>hora</i>	<i>est</i>	θ_0	P_0	z_0	UR_0	U_*	U_{H0}	lat	Δt	h_{old}	U_{1-850}	U_{2-850}	h_{g850}	U_{1-500}	U_{2-500}	h_{g500}	UR_{700}
sth6	✓				✓	✓	✓				✓				✓			
sth7	✓				✓	✓	✓		✓		✓							
sth8	✓				✓	✓	✓		✓	✓								
sth9	✓				✓	✓	✓		✓		✓							
sth10	✓		✓	✓	✓	✓	✓	✓	✓	✓								
sth11	✓		✓	✓	✓	✓	✓	✓	✓		✓							
sth12		✓	✓	✓	✓	✓	✓	✓	✓	✓								
sta1	✓		✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓
sta2	✓		✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓
sta3		✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓
sta4			✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
sta5	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
sta6			✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓
sta7					✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓
sta8			✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
sta9																		
sta1-d/sta1-n		✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓
sta2-d/sta2-n		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
sta3-d/sta3-n			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
stah1			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

3.4 Parâmetros para a avaliação de qualidade das estimativas

Os resultados obtidos pelo conjunto de validação foram submetidos a uma avaliação de qualidade, aplicando-se alguns índices estatísticos considerados representativos na literatura. Esses índices são apresentados a seguir, sendo que nas Equações de (3.2) a (3.6) E_i e O_i representam valores estimados e observados, respectivamente.

- IOA - *Index of Agreement*: índice adimensional proposto por Willmott em 1980 [27], com 1 indicando total concordância entre os valores do modelo e os valores observados.

$$IOA = 1 - \frac{\sum_{i=1}^n (E_i - O_i)^2}{\sum_{i=1}^n (|E_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (3.1)$$

- IOA_{mod} - IOA modificado: modificação do IOA, proposta por Willmott em 1985 [27].

$$IOA_{mod} = 1 - \frac{\sum_{i=1}^n |E_i - O_i|}{\sum_{i=1}^n (|E_i - \bar{O}| + |O_i - \bar{O}|)} \quad (3.2)$$

- MAE - Média dos Valores Absolutos dos Erros: dimensiona o erro das estimativas em relação ao valor observado, em módulo.

$$MAE = \frac{1}{n} \sum_{i=1}^n |E_i - O_i| \quad (3.3)$$

- MBE - Média dos Valores dos Erros (*Medium Bias Error*: é o erro médio; assim como o MAD, dimensiona o erro dos valores estimados, mas levando em consideração seu sinal, indicando tendências de super (> 0) e subestimativas (< 0)).

$$MBE = \frac{1}{n} \sum_{i=1}^n (E_i - O_i) \quad (3.4)$$

- RMSE - Raiz do Erro Médio Quadrático (*Root Mean Square Error*): é inversamente proporcional à qualidade da estimativa, ou seja, quanto maior o valor de RMSE, menor a conformidade dos resultados em relação aos valores observados.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - O_i)^2} \quad (3.5)$$

- Correlação linear (r) - é utilizado o coeficiente de correlação de Pearson, que indica o grau de correlação linear entre duas variáveis. Varia entre os valores -1 e 1, sendo que o valor 0 indica que não há correlação entre as variáveis em estudo, e 1 que há perfeita correlação.

$$r = \frac{\sum_{i=1}^n (E_i - \bar{E})(O_i - \bar{O})}{\sqrt{(\sum_{i=1}^n (E_i - \bar{E}))^2 (\sum_{i=1}^n (O_i - \bar{O}))^2}} \quad (3.6)$$

Capítulo 4

Resultados e discussão

4.1 Altura de rugosidade

Os valores obtidos para a altura de rugosidade z_0 são apresentados na Tabela 4.1.

Tabela 4.1: Resultados obtidos para z_0 para cada localidade.

Reno	Amarillo	Peachtree
0.08	0.05	0.67

Na literatura, aeroportos tem z_0 de aproximadamente 0.03 m, assim como áreas de campo com gramíneas. Já em áreas de cultivo de plantas de pequeno porte $z_0 = 0.1$ m, chegando a $z_0 = 1.0$ m para regiões com florestas bem desenvolvidas, vilas e periferia de regiões urbanas [28].

No caso de Peachtree a estação de lançamento da radiossonda se encontra no aeroporto da cidade, na região de periferia, com vegetação de grande porte no entorno. É esperado, com base na literatura, que a altura de rugosidade para essa localidade esteja entre os valores de 0.03 m e 1.0 m, com tendência aproximar-se mais de 1.0 m, devido às características da vegetação em todo o entorno do aeroporto. O valor de $z_0 = 0.67$ encontrado nesse trabalho está dentro desse intervalo.

Já em Reno e Amarillo, a vegetação é bem mais baixa e menos densa, encaixando-se nas determinações da literatura para $z_0 = 0.03$ m a $z_0 = 0.1$. No caso da cidade de Amarillo, o valor encontrado de 0.05 m foi considerado satisfatório, pois a estação está localizada em um aeroporto, em meio a uma vegetação de campo. Em Reno, a presença de alguns obstáculos como pequenas edificações da região periférica da cidade aproximam o valor de z_0 de 0.1.

4.2 Estimativa da altura da CLA pela parametrização de Panofsky

Os resultados obtidos pelo emprego da parametrização proposta por PANOFSKY [17] para camada limite instável, Equação (2.27) são apresentados e discutidos nessa seção, em comparação com as alturas estimadas pelos métodos dos gradientes e critérios subjetivos. Na Tabela 4.2, são listados h_s e h_l , representando as alturas estimadas pela parametrização considerando *lapse rate* próximo à superfície e de toda a camada, respectivamente. A altura considerada como referência para avaliação é representada como h .

Tabela 4.2: Resultados obtidos com emprego da parametrização de Panofsky.

Dia	h_s (m)	h_l (m)	h (m)
3	293	291	1276
4	198	914	1982
5	162	756	1893
6	-1015	749	1976
7	1030	423	2117
8	470	439	2284
10	-619	609	366
12	300	1284	1587
13	157	543	2265
14	198	231	209

Para o dia 11, não foram realizadas estimativas, dado que a CLA encontra-se visivelmente em condição estável e não se encaixa nos requisitos para essa parametrização.

O primeiro resultado significativo observado é que o uso do *lapse rate* superficial não é representativo da camada de mistura como um todo. Ao tomarmos os casos dos dias 6 e 10, vemos que os resultados indicam alturas negativas. Isso se deve ao fato de termos temperaturas a 10 m mais elevadas que as temperaturas do primeiro nível próximo à superfície, indicando estabilidade local. A CLA em sua totalidade apresenta *lapse rate* médio de θ_v positivo, ou seja, indicando uma situação de estabilidade estática. Para o cenário que utiliza apenas os dados de superfície, nesses casos, a parametrização não seria aplicável, visto que foi deduzida para casos instáveis.

Em 12 de junho a aplicação da parametrização com Γ_d e Λ calculados para toda a

camada resultou num valor de h_l próximo ao esperado, com erro de 20%, e um erro de 10,5% na estimativa do dia 14. Contudo, o menor erro percentual também é referente a um caso de superestimação, o que não é desejável para aplicações de dispersão.

Dois perfis são apresentados na Figura 4.1, exemplificando também os métodos utilizados para determinação da altura da CLA a partir de radiossondagens. Nessa figura, as radiossondagens são identificadas pela data e hora em UTC fornecidas diretamente pelo SPARC, e em seguida pela data e horário locais. No caso das radiossondagens de Peachtree apresentadas na Figura 4.1, os lançamentos foram realizados à 00 UTC dos dias 10 e 15 de junho. Considerando que os lançamentos são realizados meia hora antes do horário UTC de referência, os perfis são referentes às 18:30 h dos dias 9 e 14 de junho no fuso local. Na Figura 4.1, $d\theta$ e $d\theta_s$ representam os níveis determinados pelo máximo gradiente dos perfis de temperatura potencial bruta e suavizada. Da mesma forma, dq e dq_s são referentes aos níveis de mínimo gradiente para os perfis bruto e suavizado de umidade específica.

No dia 09 de junho a camada às 18 : 30 h encontra-se bem misturada e desenvolvida, como podemos ver na Figura 4.1a, e os métodos de gradiente de θ_v e q indicam uma provável camada residual.

O método do gradiente de θ_v não indica a primeira inversão de temperatura observada na radiossondagem de 13 de junho, enquanto o método do gradiente de umidade indica a redução abrupta de umidade no topo da camada limite, e se encontra no nível de 2575 m. Considerando o critério conservador, é definida para essa radiossondagem $H = 2265$ m, aproximadamente 1700 m acima do valor estimado pela parametrização $h = 543$ m.

Em termos gerais, mesmo se considerarmos as variações médias de temperatura das camadas, em lugar de aproximações com dados superficiais como suposto na dedução da parametrização, os resultados da parametrização de Panofsky não são satisfatórios, e essa formulação não se mostra representativa da dinâmica da CLA no sítio estudado.

Indo além, quando treinamos uma rede neural artificial para situações problema, em geral buscamos encontrar com essa metodologia uma relação entre os parâmetros considerados relevantes e a variável que se deseja estimar. Contudo, ao substituímos o valor real por um valor estimado por uma equação, estamos de certa forma forçando nosso modelo a “enxergar” a mesma relação já proposta pela parametrização, dado que essa é a relação exata que existe entre o valor estimado e os parâmetros de entrada. Por mais que outros dados além dos usados para obter o valor parametrizado de h sejam utilizados como entrada na rede, é possível que sejam interpretados apenas como ruído, visto que existe uma função exata que relaciona alguns dos dados de entrada ao valor h , que é o valor de referência.

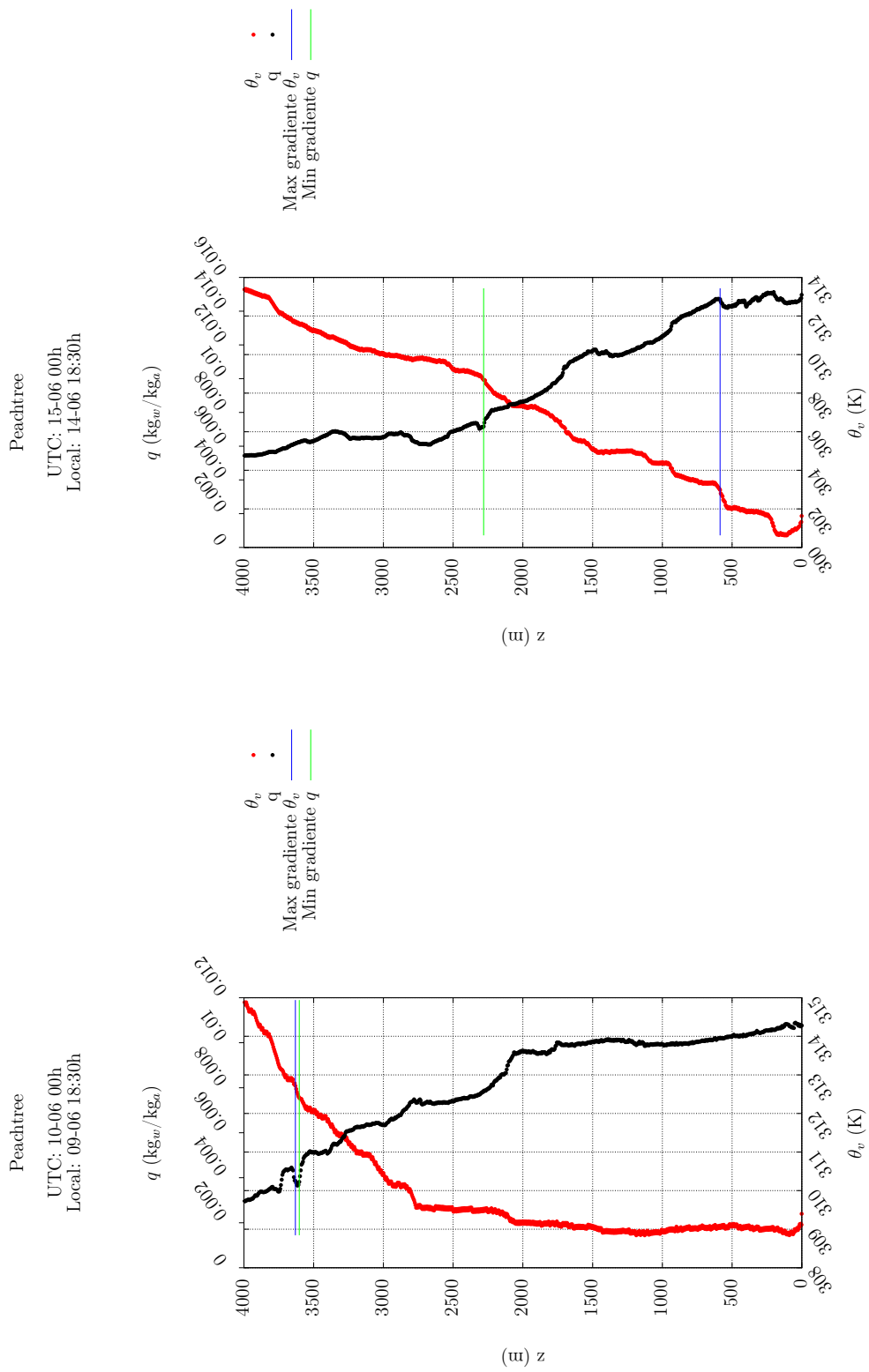


Figura 4.1: Radiossonagens da estação de Peachtree, às 18 : 30 h no horário local. (a) Radiossonagem do dia 10 de junho 00 UTC. (b) Radiossonagem do dia 15 de junho 00 UTC.

4.3 Estimativa da altura da CLA a partir de radiossondagens

Nessa sessão são apresentados alguns casos de radiossondagens para discussão. As figuras trazem os perfis de θ_v e q , e também os níveis de máximo gradiente de θ_v e mínimo gradiente de q (obtidos a partir dos perfis brutos, $d\theta$ e dq , e suavizados $d\theta_s$ e dq_s). Nessas figuras há também informações identificando qual a estação em que a radiossondagem foi realizada, bem como data e hora de acordo com os fusos UTC e local.

4.3.1 Camada de mistura - Inverno

A Figura 4.2 representa duas radiossondagens feitas no mês de janeiro na estação de Reno, Nevada. Os lançamentos foram realizados no período da tarde, e a existência de uma camada de mistura em ambos os perfis é evidente.

No caso da Figura 4.2a, o nível do valor máximo de $d\theta_v/dz$ é 823 m, 13 m abaixo do nível indicado para o valor mínimo de dq/dz . Com uma diferença de altura menor que 50 m, consideramos que os métodos concordam quanto ao nível do topo da CLA, que nesse caso é considerado como $h = 823$ m, o valor mais conservador.

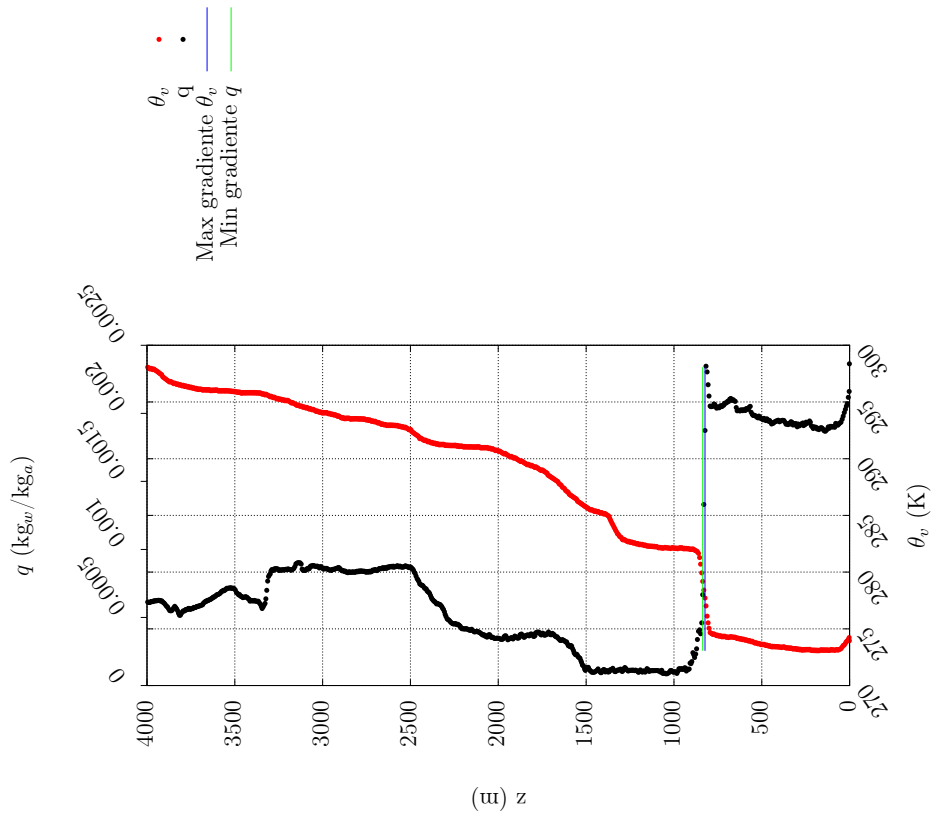
Já para a radiossondagem da Figura 4.2b, apesar de aparentemente apresentar perfis de temperatura e umidade semelhantes aos da radiossondagem da Figura 4.2a, indicando o topo da CLA a aproximadamente 1000 m, não há concordância entre os diferentes métodos aplicados. O mínimo gradiente de q encontra-se no nível de 1064 m, enquanto o máximo gradiente de θ_v está a 2505 m de altura. O perfil suavizado de θ_v tem seu máximo gradiente a uma altura de 1829 m. Para essa radiossondagem, a altura da CLA foi definida como o nível de mínimo gradiente de q , aplicando-se também um critério subjetivo, considerando as características dos perfis, principalmente o de temperatura. Essa escolha concorda também com o critério conservador, definido no início do trabalho.

Outras duas radiossondagens, apresentadas nas Figuras 4.3a de Peachtree e 4.3b de Amarillo, foram feitas no período da tarde em dias de inverno. No lançamento feito em Peachtree, assim como no apresentado na Figura 4.2a, os diferentes métodos de gradiente são coincidentes, com apenas 9 m de diferença entre eles, e $h = 729$ m.

Na radiossondagem apresentada para Amarillo na Figura 4.3b, contudo, nenhum dos métodos de gradiente automatizados foi capaz de identificar h corretamente. A altura definida subjetivamente, com base principalmente no perfil de θ_v e em sua derivada de segunda ordem, é 1021 m. Os gradientes nesse caso indicam outra inversão de temperatura e redução de umidade, em um nível mais de 1500 m superior, provavelmente referente a uma camada residual do dia anterior.

Reno

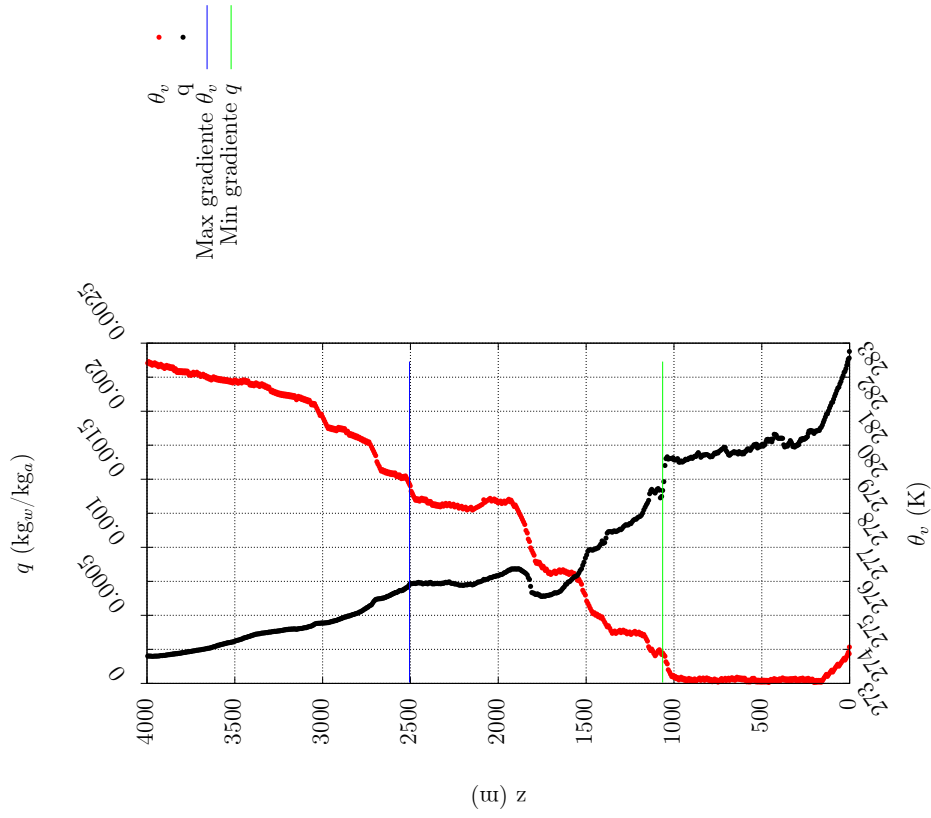
UTC: 01-01 00h
Local: 31-12 15:30h



(a)

Reno

UTC: 07-01 00h
Local: 06-01 15:30h

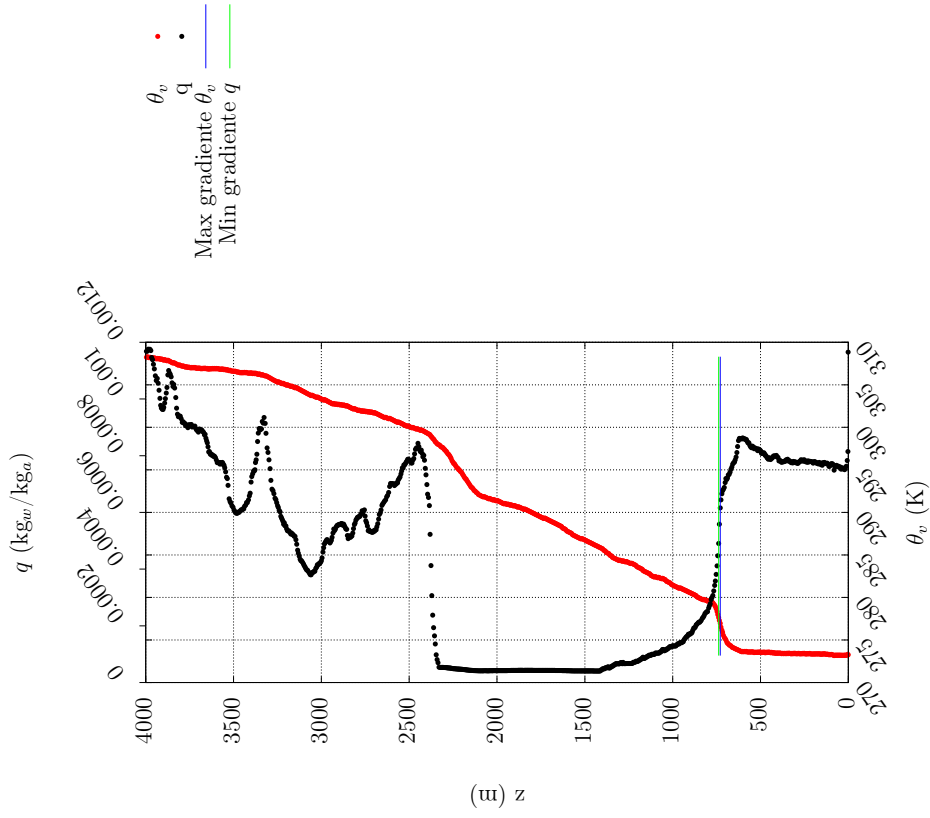


(b)

Figura 4.2: Radiosondagens da estação de Reno, com lançamento às 15 : 30 h no horário local. (a) Lançamento de 1º de janeiro, 00 UTC. (b) Lançamento de 7 de janeiro, 00 UTC.

Peachtree

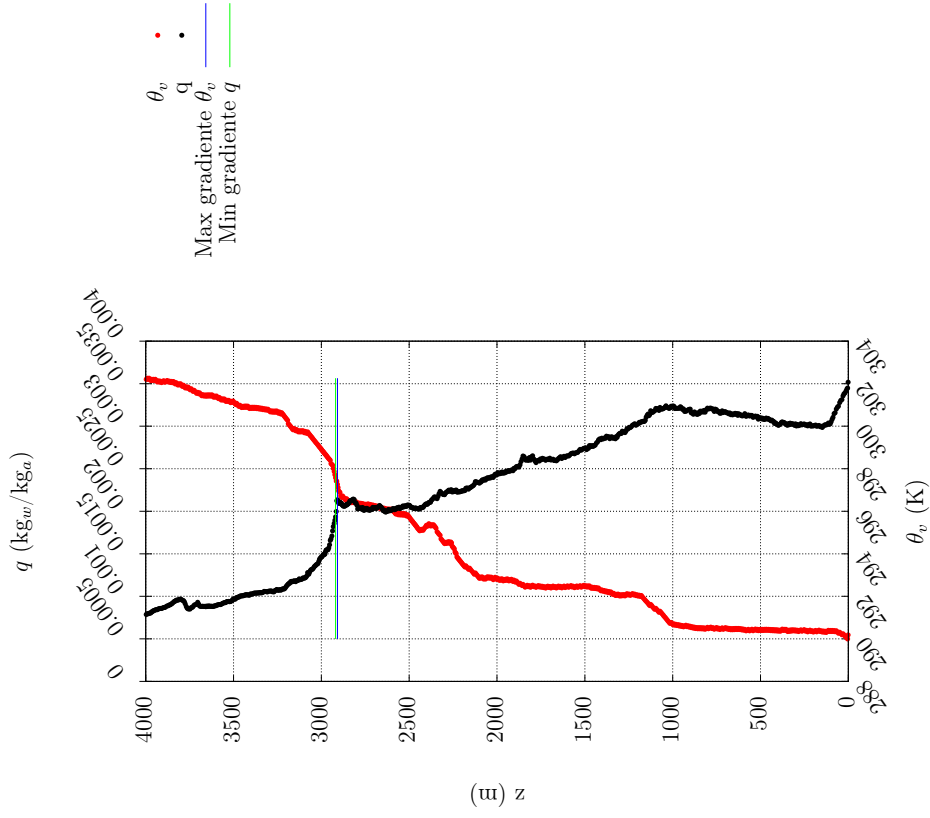
UTC: 23-12 00h
Local: 22-12 18:30h



(a)

Amarillo

UTC: 05-01 00h
Local: 04-01 17:30h



(b)

Figura 4.3: Radiosondagens com lançamento à tarde, para as estações de Amarillo e Peachtree. (a) Lançamento na estação de Peachtree, dia 23 de dezembro às 18 : 30 h no horário local, 00 UTC. (b) Lançamento na estação de Amarillo, dia 5 de dezembro às 17 : 30 h no horário local, 00 UTC.

4.3.2 Camada de mistura - Verão

Nos meses mais quentes do ano, a incidência maior de radiação solar está relacionada a camadas de mistura mais desenvolvidas ao longo da tarde.

As radiossondagens apresentadas nas Figuras 4.4a e 4.5b são das estações de Reno e Amarillo, respectivamente, e em ambas vemos que o valor mínimo de dq/dz de seus respectivos perfis indicam satisfatoriamente a altura da camada de mistura. Em 4.4a encontramos o valor de $h = 3569$ m, e para a radiossondagem de Amarillo $h = 1891$ m.

Uma segunda radiossondagem de Reno, que pode ser encontrada na Figura 4.4b, mostra uma camada também bem desenvolvida, mas há concordância entre os métodos de gradientes de temperatura e umidade. Nessa radiossondagem vemos um comportamento não convencional no perfil de umidade, o qual pode ser resultado da passagem da radiossonda por uma nuvem.

Na radiossondagem de Peachtree apresentada na Figura 4.5a, os gradientes indicam uma inversão superior a do topo da CLA, possivelmente de uma camada residual. O perfil de temperatura sugere o nível aproximado entre 1500 m e 2000 m, mas os dados calculados de derivadas de primeira e segunda ordens de temperatura não são suficientes para determinação de h . Observamos então o perfil de umidade, o qual mostra uma variação significativa em suas características, com nível passível de determinação mais precisa a partir da derivada de primeira ordem de q , já calculada. O topo da camada limite para essa radiossondagem foi definido então como $h = 1860$ m.

4.3.3 Camada noturna

A estimativa da camada limite noturna se dá de forma diferente da diurna. Em geral, vemos uma inversão de temperatura com base na superfície, e h pode ser definido então como o nível do topo da inversão, e não mais de base. Uma outra maneira é identificar o primeiro máximo de u_H , ou jato de baixa altitude [13].

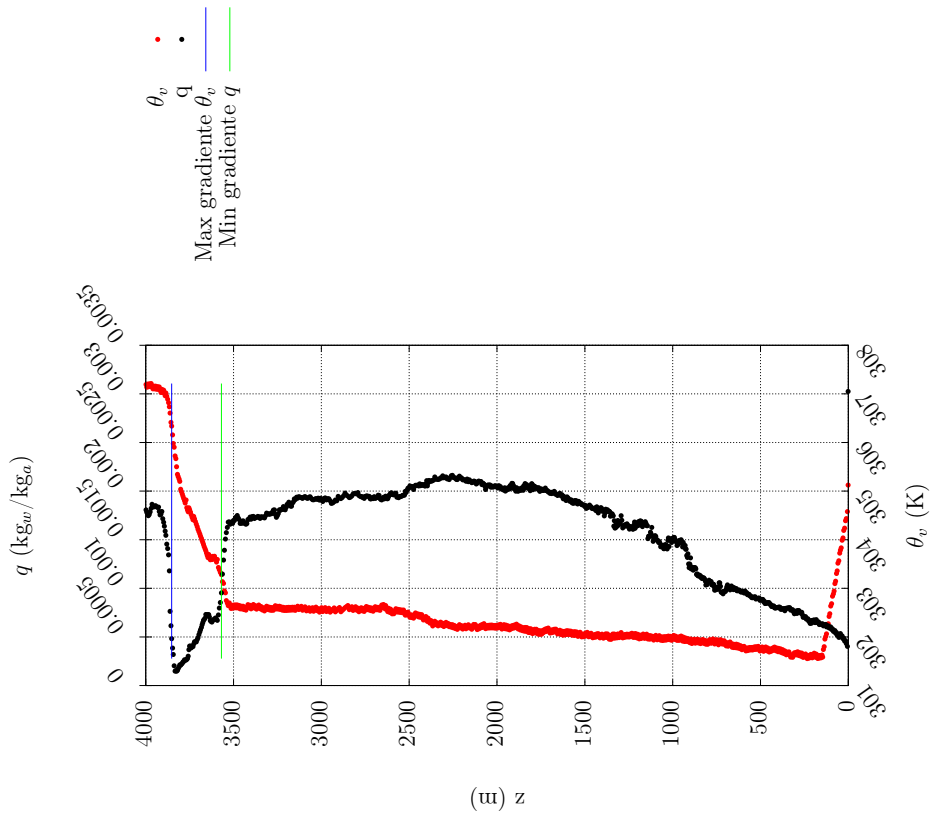
No caso da radiossondagem apresentada na Figura 4.6, não pudemos identificar um *LLJ* (jato de baixo nível), e o gradiente de temperatura indica um primeiro nível possível, mas se comparado com o perfil de umidade parece subestimar h , e talvez indique uma subcamada. A altura da CLA foi definida pelo gradiente de umidade como $h = 229$ m.

A altura da camada limite na radiossondagem de Amarillo, representada na Figura 4.7, foi definida como $h = 111$ m, utilizando o critério de *LLJ*. A variação abrupta de umidade e a mudança na inclinação do perfil de temperatura próximas ao nível de 2000 m são indicativas de uma possível camada residual.

Na estação de Peachtree, com lançamentos às 6:30 h (12 UTC), as manhãs do período de verão são marcadas pela presença de um jato de baixo nível bastante característico,

Reno

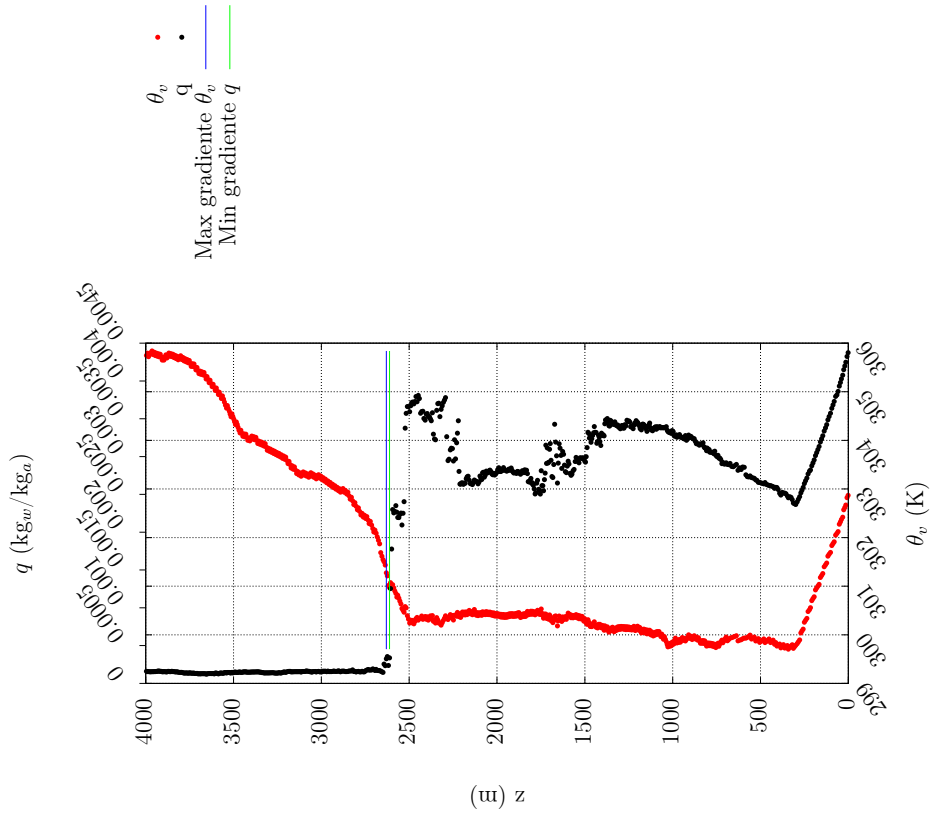
UTC: 02-07 00h
Local: 01-07 15:30h



(a)

Reno

UTC: 10-06 00h
Local: 09-06 15:30h

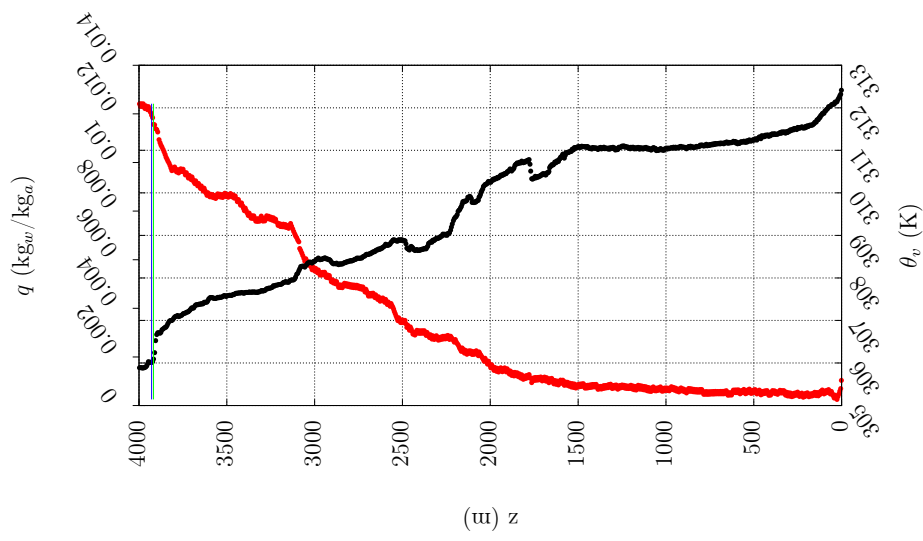


(b)

Figura 4.4: Radiossonagens com lançamento à tarde, meses de junho e julho. (a) Radiossonagem da estação de Reno, com lançamento às 15 : 30 h do dia 1º de julho no horário local, 00 UTC. (b) Radiossonagem da estação de Reno, com lançamento às 15 : 30 h do dia 10 de junho no horário local, 00 UTC.

Peachtree

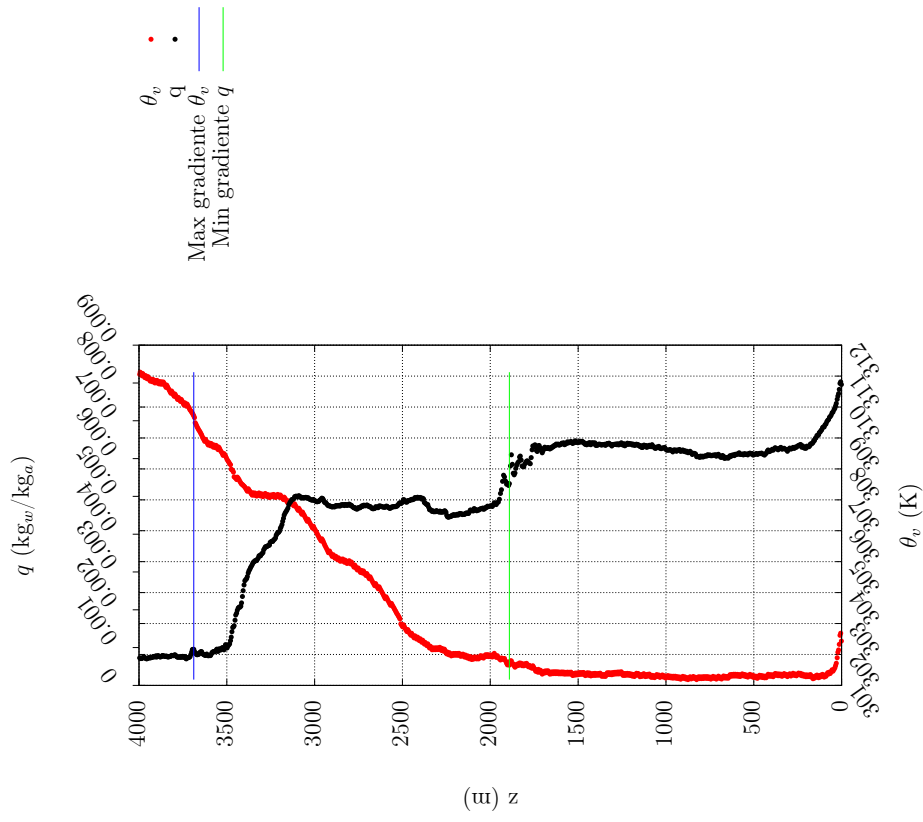
UTC: 23-06 00h
Local: 22-06 18:30h



(a)

Amarillo

UTC: 30-06 00h
Local: 29-06 17:30h



(b)

Figura 4.5: Radiossondagens com lançamento à tarde. (a) Radiossondagem na cidade de Peachtree, às 18 : 30 h no horário local, 00 UTC. (b) Radiossondagem na cidade de Amarillo, às 17 : 30 h no horário local, 00 UTC.

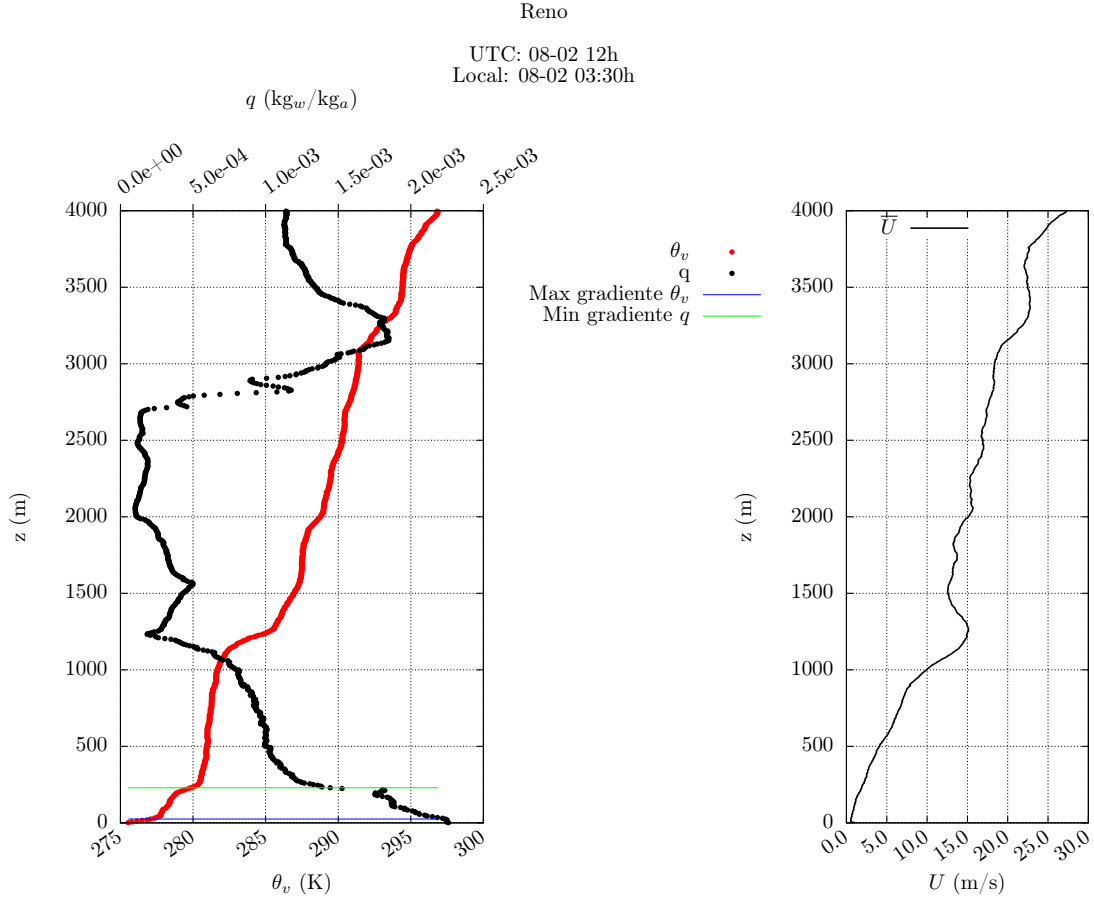


Figura 4.6: Lançamento às 3:30 h para a estação de Reno, fevereiro.

como é possível observar na Figura 4.8, utilizado para encontrar a altura da CLA, que nesse caso é de 313 m.

Uma segunda radiossondagem de Peachtree, Figura 4.9, apresenta $h = 305$ m, definida subjetivamente por observação dos perfis e de suas derivadas.

4.3.4 Dificuldades técnicas

Os parâmetros estimados L , u_* e θ_* foram obtidos para cada radiossondagem, empregando as Equações (2.7), (2.8) e (2.9) de forma iterativa. Em alguns casos não houve convergência, diagnosticada como resultado de casos com z/L fora dos limites para os quais a teoria de Monin-Obukhov é válida.

Além disso, vários casos não apresentam características conclusivas para determinação de h . Dois exemplos são apresentados nas Figuras 4.10 e 4.11. Nessas radiossondagens não há a presença de um jato de baixo nível característico para determinação da camada noturna. Tampouco há indicativos claros nos perfis de umidade e temperatura, sendo descartados para esse trabalho.

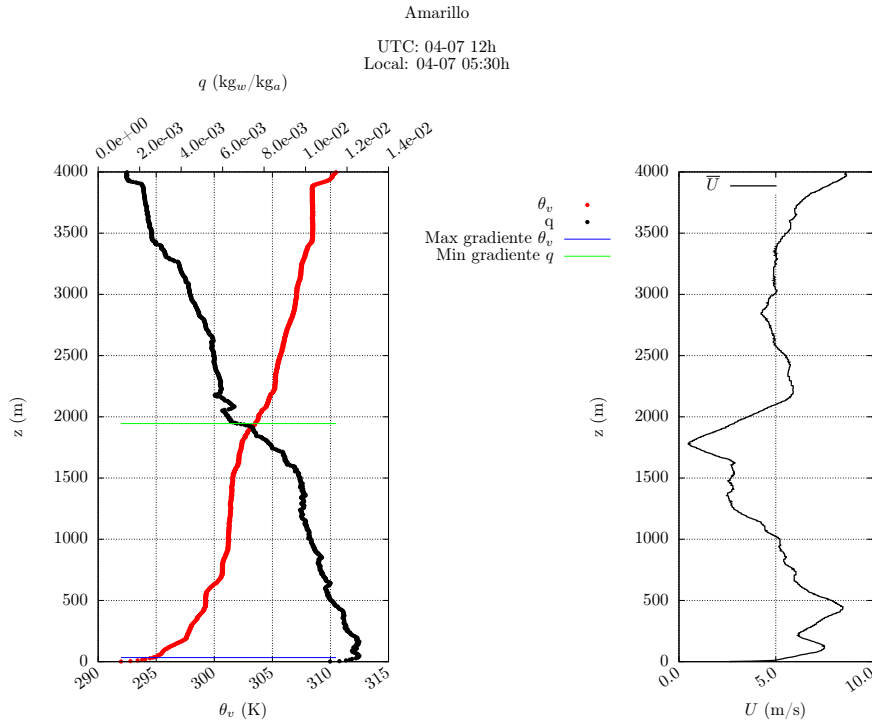


Figura 4.7: Lançamento às 5:30 h para a estação de Amarillo.

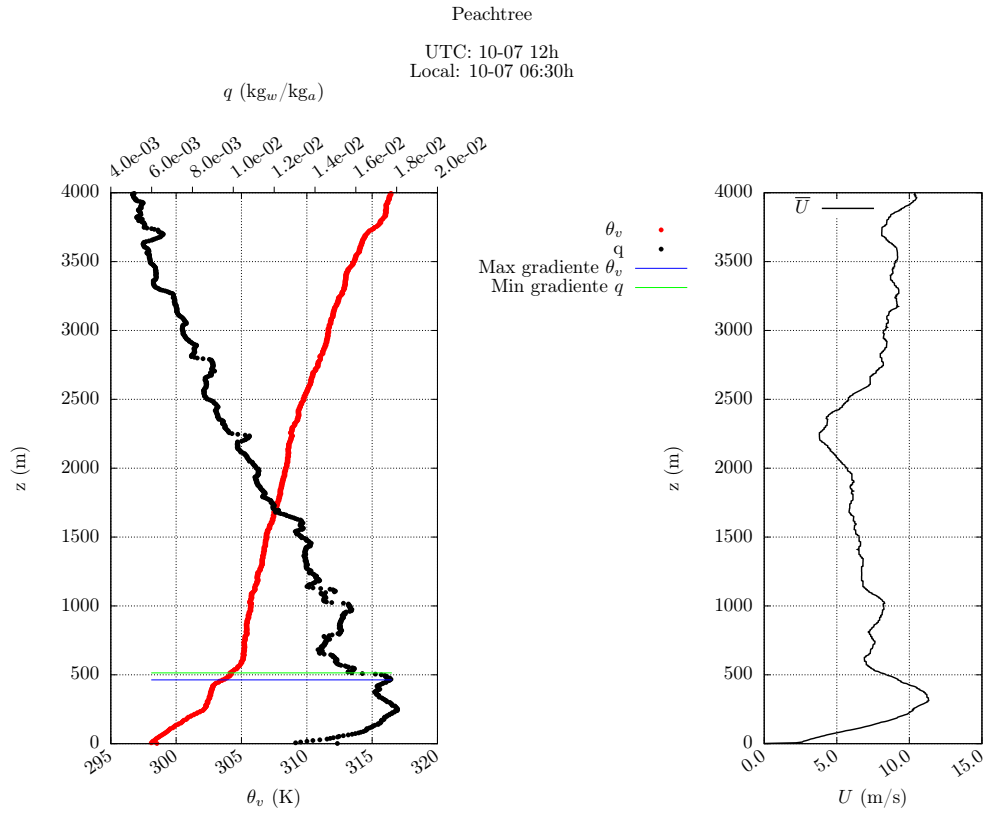


Figura 4.8: Lançamento às 6:30 h para a estação de Peachtree.

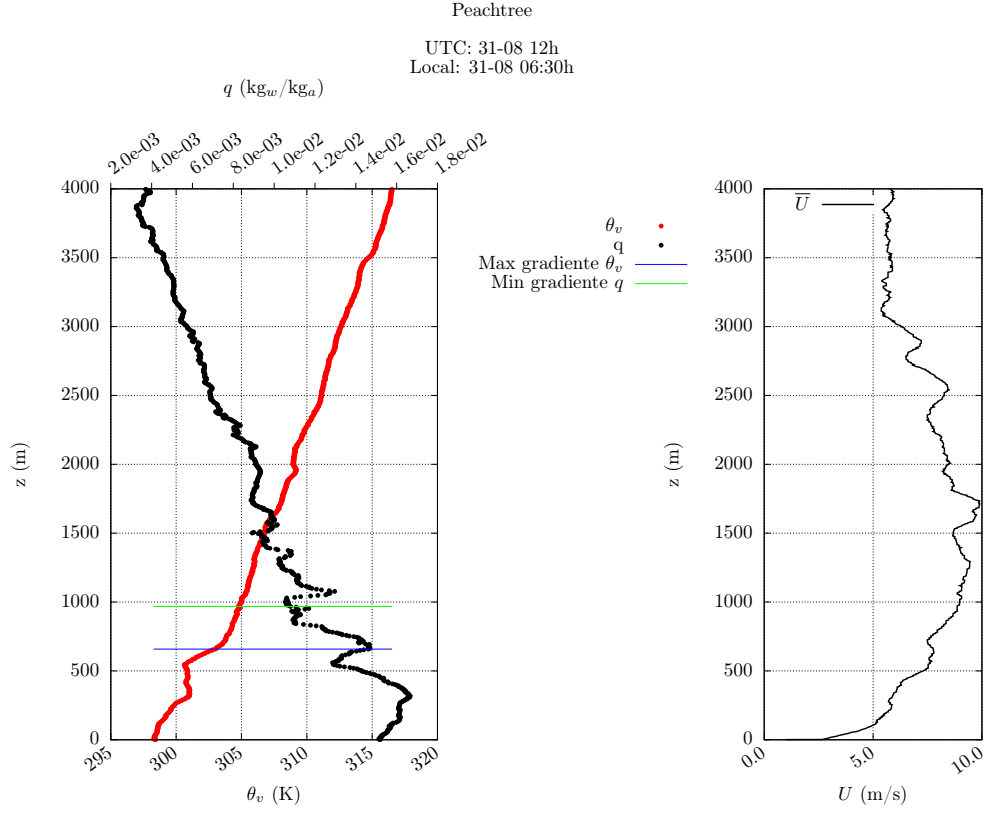


Figura 4.9: Lançamento às 6:30 h para a estação de Peachtree.

Os dados de altura da CLA estimados e que poderão ser utilizados para aplicação na RNA totalizam 486 registros, sendo 380 registros para Reno e 106 registros para Peachtree, e são disponibilizados no Apêndice A, junto aos parâmetros estimados com a aplicação da teoria de Monin-Obukhov. Aproximadamente 50 registros de estimativa de h para Reno e 4 registros para Peachtree foram descartados, por não ter sido possível a estimativa de L , u_* e θ_* para esses casos.

4.4 Estimativa da altura da CLA com emprego da RNA

Nesta seção são apresentados alguns resultados obtidos com emprego da rede neural artificial para estimativa de h . Foram realizados diversos experimentos, considerando diferentes configurações. Além disso, após o treinamento de uma única rede generalizada testamos também o emprego de duas redes especializadas, uma para casos de CLA diurna e outra para casos de CLA noturna.

Primeiramente, são apresentados resultados obtidos ao empregar-se uma única rede generalizada, separando-os por bloco de experimentos (st , sth e sta). Em seguida, alguns

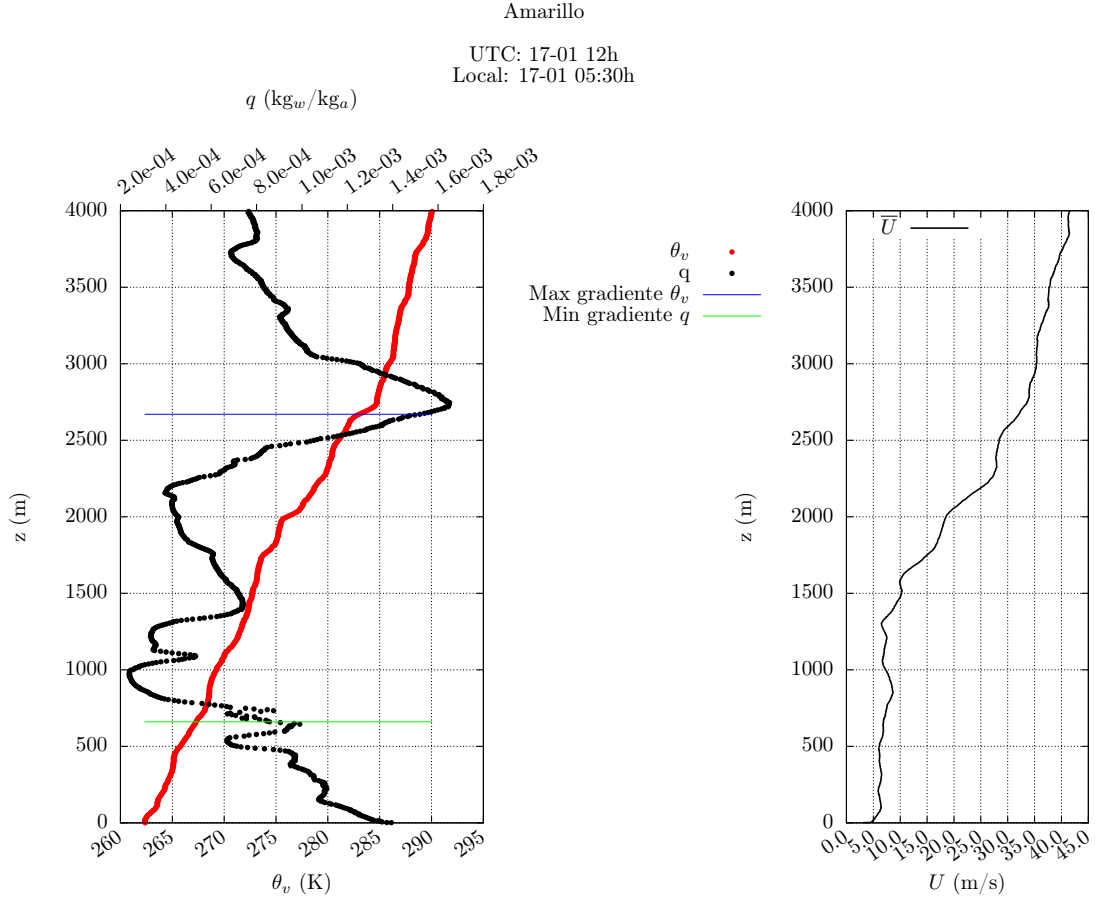


Figura 4.10: Lançamento às 05:30 h, horário local, para a estação de Amarillo.

resultados encontrados ao empregarmos duas redes especializadas são apresentados.

Os resultados de todas as configurações estão disponíveis no Apêndice A.

4.4.1 Estimativa de h com emprego de uma RNA generalizada

Bloco de experimentos *st*

Neste bloco de experimentos foram utilizados como dados de entrada:

- Parâmetros meteorológicos de superfície (θ_0 , P_0 , UR_0 , U_{H0} e u_*),
- Parâmetros para caracterização do sítio (lat e z_0),
- Parâmetros para determinação dos períodos do dia ($hora$) e do ano (est).

A primeira configuração testada para esse bloco, *st1*, incluiu todos os parâmetros citados acima. Nas configurações seguintes, cada um dos parâmetros de entrada foi retirado isoladamente, e também em conjunto com outros parâmetros.

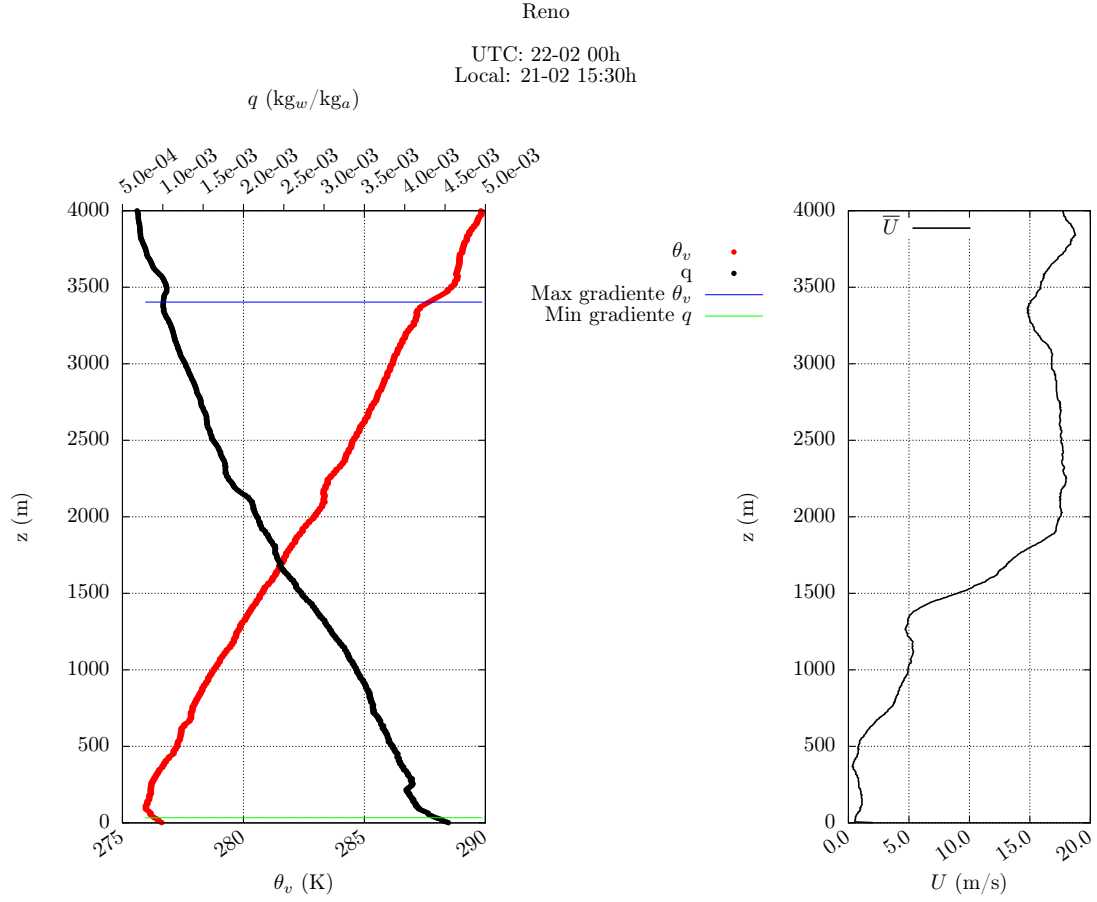


Figura 4.11: Lançamento às 15:30 h, horário local, para a estação de Reno.

Utilizando como critério para determinação da melhor configuração o IOA, são apresentados os resultados referentes às três melhores configurações do bloco *st* na Tabela 4.3. Os maiores valores de IOA e r correspondem à configuração *st4*, cujo gráfico de dispersão é apresentado na Figura 4.12, e na qual u_* é retirado do vetor de entrada. Contudo, *st7* atingiu o maior valor de IOA_{mod}, e também o menor valor de MAE, e o parâmetro ausente nessa configuração é a pressão no nível da superfície (P_0).

Tabela 4.3: Resultados obtidos para as melhores configurações do bloco de experimentos *st*.

Configuração	IOA	IOA _{mod}	r	RMSE	MAE	MBE
st4	0.9432	0.8113	0.9017	426.4814	301.5864	-14.4146
st6	0.9418	0.8094	0.8975	427.0978	303.1855	-11.1716
st7	0.9417	0.8158	0.8988	429.8743	297.1245	-24.2191
st7-ln	0.9265	0.7960	0.8694	486.8685	330.9670	-59.1836

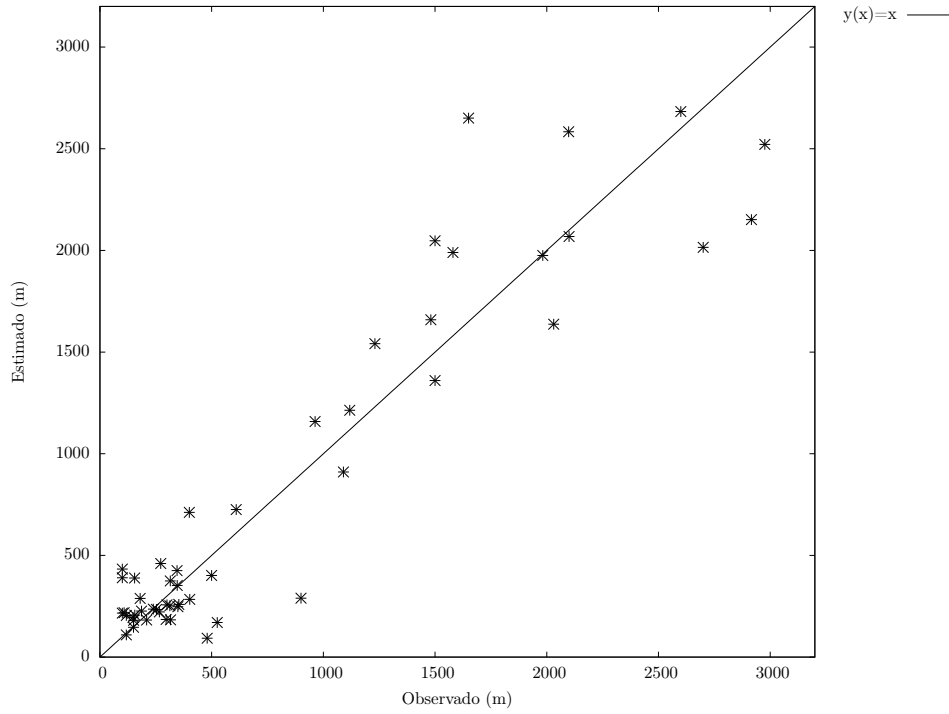


Figura 4.12: Gráfico de dispersão para a configuração st4.

A Tabela 4.3 traz também os resultados referentes ao experimento st7-ln, com configuração de parâmetros de entrada semelhante ao experimento st7, mas com o valor de h substituído por seu logarítimo natural ($\ln(h)$).

Os menores valores de IOA foram encontrados para as configurações st11 e st14, e podem ser encontrados na Tabela 4.4 juntamente aos outros índices estatísticos calculados para esses experimentos. Ambas as configurações não incluíram os parâmetros *hora* e *est* como entrada da RNA, diferindo entre si apenas pela inclusão da *lat* na configuração st14.

Tabela 4.4: Resultados obtidos para as piores configurações do bloco de experimentos *st*.

Configuração	IOA	IOA _{mod}	r	RMSE	MAE	MBE
st11	0.8873	0.7060	0.8097	559.3019	424.4469	34.3531
st14	0.8926	0.7195	0.8217	523.1814	404.0557	-1.3187

Bloco de experimentos *sth*

Neste bloco de testes são considerados, além dos parâmetros já incluídos no bloco st, as informações de h_{old} e Δt . Mesmo os melhores resultados desse bloco, disponíveis na Tabela 4.5, atingiram valores menores de IOA que o bloco st, sem a inclusão de h_{old} e Δt , e um gráfico de dispersão da configuração sth10 pode ser encontrado na Figura 4.13.

Tabela 4.5: Resultados obtidos para as melhores configurações do bloco de experimentos *sth*.

Configuração	IOA	IOA _{mod}	r	RMSE	MAE	MBE
sth4	0.9391	0.7998	0.8940	429.0692	311.4511	20.6295
sth10	0.9335	0.7962	0.8822	461.5054	326.1548	11.0901

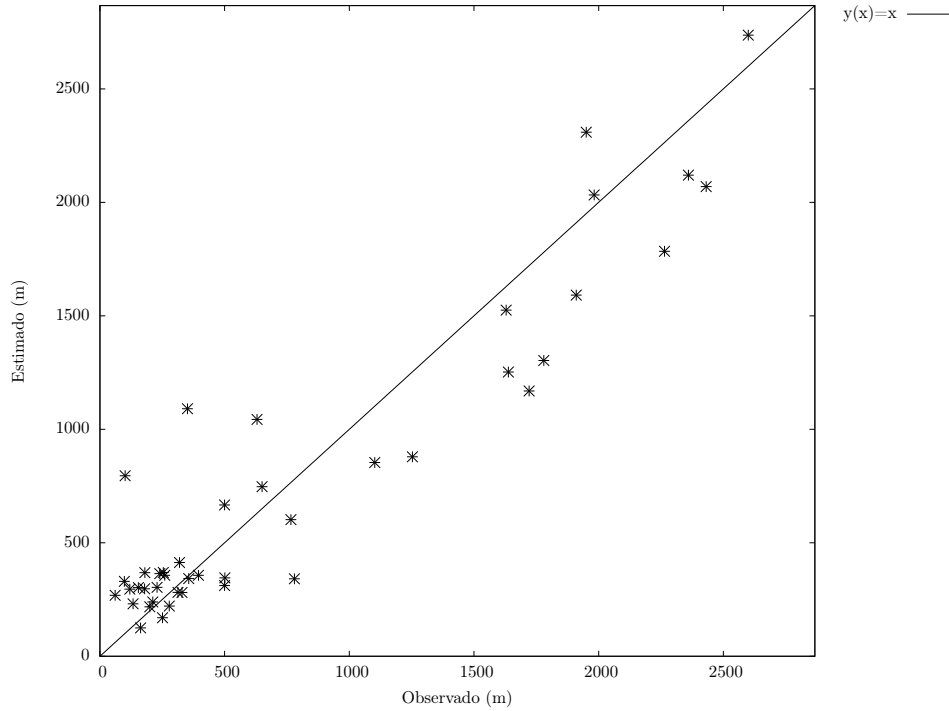


Figura 4.13: Gráfico de dispersão para a configuração sth10.

A configuração deste bloco com piores resultados foi a sth2, a qual inclui em seu vetor de entrada apenas os parâmetros h_{old} e $lapse$, com $IOA = 0.5776$ e $r = 0.3181$.S

Bloco de experimentos *sta*

O bloco de experimentos apresenta a inclusão dos parâmetros de altitude para representação de fenômenos de escala sinótica. Em um dos testes do bloco, o sta8, foi avaliada também a inclusão de h_{old} e Δt .

A configuração sta8 apresentou resultados piores que os obtidos apenas com dados de superfície, com uma tendência a subestimar os valores, o que pode ser observado na Figura 4.14, e também pelo valor negativo de $MBE = -52.7677$.

Com seus resultados disponibilizados na Tabela 4.6, as configurações sta5 e sta2 diferem pela inclusão ou não, respectivamente, da latitude das estações, bem como do período do ano. A configuração sta2 apresentou os maiores valores de IOA, IOA_{mod} e r, além do

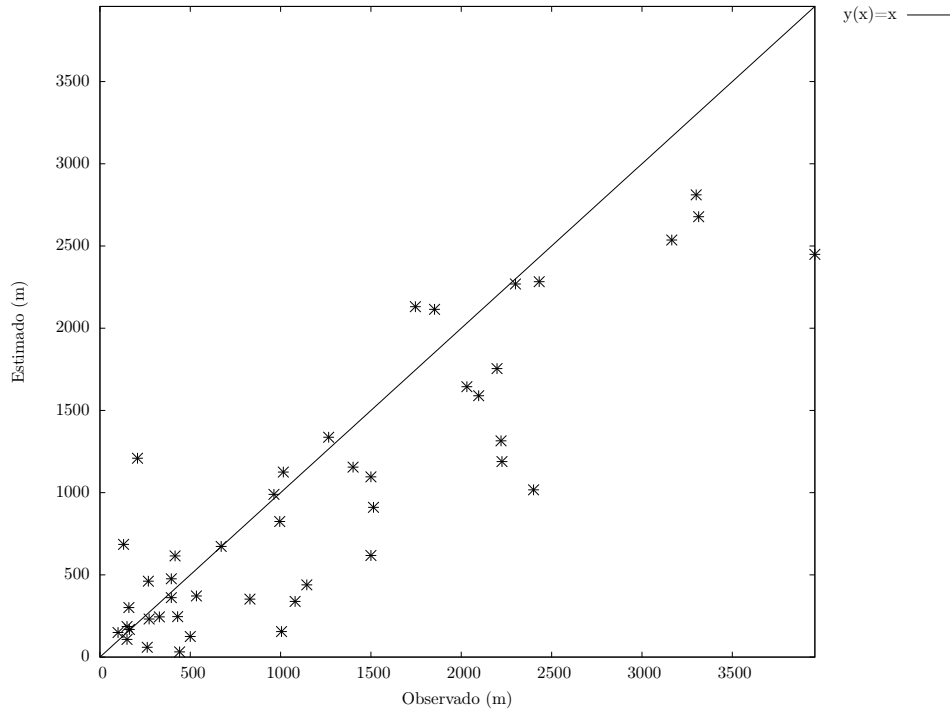


Figura 4.14: Gráfico de dispersão para a configuração sta8.

menor valor de MAE. O vetor de entrada da segunda melhor configuração desse bloco, sta1, se assemelha ao de sta2, exceto pela inclusão do parâmetro *est*. Na Figura 4.15 é representado um gráfico de dispersão para a configuração sta5.

Tabela 4.6: Resultados obtidos para as melhores configurações do bloco de experimentos *sta*.

Configuração	IOA	IOA _{mod}	r	RMSE	MAE	MBE
sta1	0.9333	0.8023	0.8814	463.3715	322.6634	35.8751
sta2	0.9473	0.8194	0.9072	407.4014	288.8027	-19.3970
sta5	0.9331	0.7935	0.8853	446.2744	321.3631	-2.4894

4.4.2 Estimativa de h com emprego de duas RNA's especializadas

Em lugar de inserirmos a informação *hora*, os pares de treinamento foram divididos de acordo com os dois períodos, representando dia e noite. Alguns resultados dessas configurações são apresentados na Tabela 4.7. Os resultados para o período noturno são invariavelmente piores que os resultados para o período diurno.

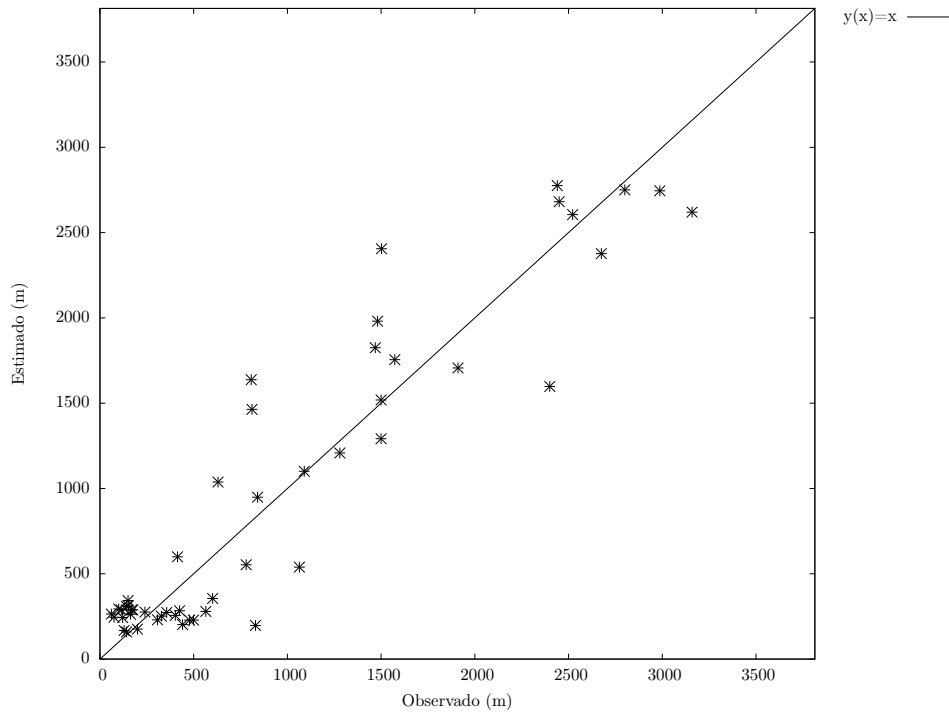


Figura 4.15: Gráfico de dispersão para a configuração sta5.

Tabela 4.7: Resultados obtidos com a separação dos dados do dia e da noite.

Configuração	IOA	IOA _{mod}	r	RMSE	MAE	MBE
dsa3	0.8034	0.5925	0.6701	656.9168	520.1658	32.4322
nsa3	0.4780	0.3869	0.2522	223.0556	146.9516	-2.1526
ds4	0.8766	0.6824	0.8005	518.3684	412.3393	46.1580
ns4	0.5425	0.4376	0.3903	211.4722	143.7493	-12.0328
dsta2	0.8327	0.6415	0.7217	624.5283	490.1128	80.0897
nsta2	0.4180	0.3331	0.1861	199.3452	142.8677	38.4041

4.5 Discussão

Os resultados que obtivemos com o emprego da rede neural mostram o alto potencial dessa ferramenta na resolução de problemas não lineares de alta complexidade. Além disso, a característica não linear do problema fica evidenciada pelos resultados, nos experimentos em que cada parâmetro é avaliado individualmente.

A mesma metodologia de avaliar o peso de cada parâmetro foi apresentada por LI [13]. Contudo, o interesse do estudo era diferente, e o objetivo do autor foi encontrar uma formulação linear para determinação de h . Quando temos problemas com características

lineares, a contribuição de cada variável se dá de forma diferente, e ao observarmos uma piora dos índices estatísticos com a retirada de um dos parâmetros do vetor de entrada, sabemos que independente dos outros parâmetros envolvidos, a retirada daquela variável em questão irá sempre piorar os resultados. Os dados de h utilizados por LI [13] foram obtidos através de parametrizações em um projeto de reanálise, e as relações entre os parâmetros de entrada não são necessariamente correspondentes às relações observadas em situações reais.

No presente estudo, fica clara a não linearidade do problema. No bloco de experimentos st, todos os parâmetros considerados de superfície, período do dia e do ano, e caracterização dos sítios foram inseridos na configuração st1, que obteve $IOA = 0.9406$, e $r = 0.8956$. Nas outras configurações, a retirada, separadamente, dos parâmetros P_0 e z_0 , respectivamente, resultou numa melhora dos índices. Na configuração st10 foram retirados os dois parâmetros conjuntamente, mas houve uma piora do desempenho da rede, com $IOA = 0.9373$ e $r = 0.8902$, o que não seria esperado em um problema com relações lineares entre as variáveis.

Já o trabalho de PELLICIONI [25] tem o mesmo objetivo ao qual nos propusemos nesse estudo, a estimativa de h a partir de dados de estações meteorológicas. Contudo, a base de dados de PELLICIONI é de valores horários de h , e foi obtido com o emprego de parametrizações, podendo não ser representativa das relações reais observadas. Como já mencionado anteriormente, uma rede neural artificial é um aproximador de funções, e com o emprego de um vetor de saída composto por valores parametrizados é esperado que a rede encontre apenas as relações entre variáveis impostas pela parametrização. A melhor configuração de rede no trabalho de PELLICIONI atingiu um valor de correlação linear de $r = 0.80$, e não são informados outros índices para comparação. A maior parte das configurações testadas no presente estudo teve um desempenho mais satisfatório, com valores de $r \geq 0.85$, como pode ser observado no Apêndice A

Os trabalhos que avaliam a determinação de h com uso de equipamentos de sensoria-mento remoto baseado no solo são numerosos, e alguns trazem outros índices de avaliação de qualidade de suas medições.

O estudo comparativo de HENNEMUTH e LAMMERT [2] entre alturas da CLA determinadas por LIDAR e por métodos clássicos a partir de radiossondagens, traz um valor de $MAD = 400$ m para o método da parcela, não sendo informados parâmetros para comparação com os métodos de gradiente. Esse é um parâmetro de difícil emprego na comparação entre trabalhos que utilizam amostras diferentes. Se considerarmos os valores determinados a partir das radiossondagens como referência, vemos que h varia entre 500 m e 3500 m, aproximadamente. Contudo, não temos informações sobre a distribuição desses valores, tampouco qual seja o valor de \bar{h} . Com valor máximo de H de 3957 m, o

presente estudo obteve valores diversos para o índice MAD. Com muitos casos com valores próximos a 500 m, configurações com melhor desempenho atingiram valores entre 300 m e 350 m.

Algo que deve ser levado em consideração quando comparamos nossos resultados aos obtidos por equipamentos de LIDAR e SODAR é a diferença nos fundamentos dos métodos. Enquanto equipamentos de sensoriamento remoto fazem medições de perfis, assim como radiossondas, os dados fornecidos à RNA como entrada contém informações apenas de superfície e caracterização das estações. Mesmo os dados de altitude são limitados, com informações de apenas 3 níveis de altitude, os quais não representam a atmosfera no momento da estimativa, mas sim em algum momento pretérito com diferença temporal usualmente entre 12 h e 24 h.

No caso das medições de h com emprego de SODAR apresentadas por BEYRICH [1], é realizada uma avaliação subjetiva dos perfis obtidos a partir do equipamento, e subsequente determinação da altura. A aplicação de métodos automatizados não é recomendada pelo autor.

Voltando às determinações a partir de perfis medidos por LIDAR, é importante ressaltar que as medições instantâneas tem um intervalo de alturas de 350 m previamente determinado, a partir de um valor de \bar{h} estimado por um segundo método a partir do LIDAR, para o período no qual se pretende estimar h , usualmente de 30 min. Não foram apresentados resultados referentes às medições instantâneas sem essa informação prévia. No caso da RNA, não há limites para a estimativa da rede, e ainda assim os resultados obtidos não se mostram muito inferiores aos do LIDAR.

Quando optamos pela rede especializada, dividindo o grupo de pares em dois subgrupos de acordo com o período do dia, pretendemos separar pares que provavelmente tem relações bastante diversas entre seus vetores de entrada e saída. O que pudemos observar já era esperado: com dois grupos aproximadamente do mesmo tamanho, os resultados para camada no período da madrugada e início da manhã teve resultados significativamente piores. A melhor configuração para o período noturno alcançou $IOA = 0.54$, e $r = 0.39$. Já os resultados do período diurno apresentaram ligeira piora em relação ao uso de todos os pares unificadamente. Isso se deve, provavelmente, a diminuição do número de pares entrada-saída para o treinamento da rede, o qual foi reduzido pela metade.

Dentre os experimentos com emprego da RNA generalizada, a configuração que obteve os melhores resultados para os índices quase todos os índices considerados nesse trabalho foi a sta2, seguida pela st4 e sth4. Os índices calculados para essas configurações são reapresentados a seguir, na Tabela 4.8, para comparação.

Enquanto sth4 apresenta tendência a superestimar os valores de h , os dois melhores resultados, sta2 e st4 possuem valores de MBE negativos, indicando uma subestimação.

Tabela 4.8: Resultados obtidos para as melhores configurações com emprego da RNA generalizada.

Configuração	IOA	IOA _{mod}	r	RMSE	MAE	MBE
sta2	0.9473	0.8194	0.9072	407.4014	288.8027	-19.3970
st4	0.9432	0.8113	0.9017	426.4814	301.5864	-14.4146
sth4	0.9391	0.7998	0.8940	429.0692	311.4511	20.6295

Se considerarmos para referência os valores de IOA e r , podemos dizer que sta2 e st4 obtiveram resultados sem diferenças significativas, e uma operação utilizando como entrada apenas parâmetros de estações meteorológicas de superfície se equipara a uma que considere também parâmetros de altitude. Isso facilitaria a operação, visto que não é necessário o lançamento de radiossondas para obtenção de parâmetros de altitude, tampouco o emprego desses parâmetros obtidos através de modelos.

Como já mencionado previamente, foi necessário o desenvolvimento de uma base de dados para treinamento e validação da RNA. Tal etapa foi certamente a mais difícil e trabalhosa de todo o estudo. Para garantir a qualidade dos dados, a estimativa de h a partir dos perfis de θ_v e q foi realizada subjetivamente, o que limitou o número de pares entrada-saída. Além disso, a disponibilização de dados de radiossondagens de alta resolução diários foi encontrada apenas para estações nos EUA, limitando nossa base de dados geograficamente.

Uma das consequências da limitação geográfica é o horário de lançamento das radiossondas. Os dados utilizados para treinamento da rede são referentes aos horários de 12 e 00 UTC. Considerando o horário local das três estações escolhidas para esse trabalho temos dados referentes a 4 h e 16 h para Reno, 6 h e 18 h para Amarillo e 7 h e 19 h para Peachtree. No intervalo entre as 7 h da manhã e as 16 h no período da tarde não temos informações de radiossondagem que possibilitem a estimativa subjetiva de h , e sua posterior inclusão para o treinamento da rede. Durante esse intervalo é comum observarmos o desenvolvimento da CLA, em resposta à presença de radiação solar, passando de uma camada fina e usualmente em condições estáveis-neutras para uma camada de mistura bem desenvolvida.

Capítulo 5

Conclusão

No presente trabalho, foi estudada a aplicação de uma rede neural artificial do tipo *multilayer perceptron* para a determinação da altura da camada limite atmosférica, um parâmetro meteorológico de extrema importância em modelos de dispersão atmosférica. O objetivo principal foi o emprego apenas de dados que possam ser obtidos em estações meteorológicas, como dados de superfície e algumas informações de altitude retiradas de radiossondagens anteriores ao momento em que se deseja estimar h .

Foi necessária a criação de uma base de dados contendo um número suficiente de medições de h confiáveis a serem utilizadas como dados observados no treinamento e validação da rede. Para tanto, perfis de θ_v e q obtidos de radiossondagens de alta resolução foram observados, considerando os níveis de máximo gradiente de θ_v e mínimo gradiente q para determinação da altura da camada limite para cada radiossondagem. Ao longo desse processo ficou clara a dificuldade de criar metodologias que estimem h de forma automatizada, pois mesmo a observação de uma ou mais pessoas com conhecimentos técnicos na área algumas vezes não era o suficiente para determinar h a partir dos perfis.

Nenhum método encontrado na literatura teve desempenho considerado satisfatório para determinação automatizada de h , sendo que o LIDAR apresentou-se como a melhor das alternativas estudadas para camada limite convectiva, e na ausência de nuvens, pois toda medição acima do nível da base das nuvens é considerada não confiável. O uso do LIDAR, contudo, na estimativa instantânea de h , como é feito quando determinamos h a partir dos perfis de radiossondagem, não pode ser avaliada, pois as condições apresentadas não foram consideradas suficientes para tal.

O uso do LIDAR para CLA noturna também não é indicado, pois nesses casos h muitas vezes encontra-se abaixo do limite inferior de detecção do equipamento. Para camadas limite muito finas uma solução proposta na literatura é a utilização de um *mini-SODAR*, mas apenas com a avaliação subjetiva dos resultados do equipamento por um técnico experiente.

Dessa forma, decidimos utilizar como dado de entrada apenas as alturas que puderam ser estimadas a partir da radiossondagem, após avaliação subjetiva dos perfis, a fim de dar mais confiabilidade ao banco de dados, e não embutir erros no treinamento.

Para operação da RNA, foram determinados 65 diferentes experimentos, cada qual com uma diferente combinação de parâmetros em seu vetor de entrada. As configurações foram divididas em três blocos de experimentos, de acordo com as características dos parâmetros de entrada:

- *st*: parâmetros de estações meteorológicas de superfície, hora, mês e caracterização do sítio;
- *sth*: parâmetros de estações meteorológicas de superfície, hora, mês, caracterização do sítio e o valor da última altura da CLA medida antes do momento em que se deseja estimar h , bem como o intervalo temporal decorrido;
- *sta*: parâmetros de estações meteorológicas de superfície, hora, mês, caracterização do sítio, parâmetros de altitude, valor da última altura da CLA medida antes do momento em que se deseja estimar h e o intervalo temporal decorrido.

Dentre tais experimentos, estão incluídos não só aqueles nos quais se empregou uma rede generalizada, como também experimentos para os quais os pares entrada-saída foram divididos entre "noturnos" e "diurnos", a fim de treinar duas redes especializadas.

Os resultados da operação da rede mostraram claramente a característica não linear das relações entre as variáveis, e a importância de cada parâmetro de entrada não se torna clara, salvo algumas exceções, como a informação de período do dia, que melhorou os resultados sempre que incluída, independentemente dos outros parâmetros considerados nas configurações.

Para os casos de rede generalizada, as melhores configurações foram *sta2*, com $IOA = 0.9473$, e *st4*, com $IOA = 0.9432$. A diferença entre os valores de IOA encontrados para *sta2* e *sta4* é de apenas 0.0041, mostrando que a inclusão de dados de altitude não resulta em melhora significativa dos resultados, quando comparada ao uso apenas de dados de estações meteorológicas de superfície e informações para caracterização dos sítios e períodos do dia e do ano. Assim, podemos concluir que há grande potencial na operação da rede neural artificial para determinação de h , sendo uma alternativa a ser considerada principalmente em quando não há disponibilidade de radiossondagens, visto que não é necessária a utilização de dados de altitude.

A utilização de duas redes especializadas por horário obteve resultados inferiores aos da rede generalizada. Para o período diurno, os índices apresentaram uma ligeira piora. Já para o período noturno, a melhor configuração encontrada alcançou um valor de

$IOA < 0.6$. A redução do número de pares entrada-saída para treinamento é uma possível justificativa para piora dos resultados, se comparados com a rede generalizada. Assim, com base nos resultados obtidos nesse trabalho, não é indicada a substituição de uma rede generalizada por duas redes especializadas por horário. Contudo, novos experimentos com uma base de dados composta por um número maior de registros de h , podem apresentar conclusões diferentes.

A comparação dos resultados obtidos com outros resultados da literatura não é clara, por dois motivos principais:

- Nenhum dos trabalhos encontrados na literatura se assemelha ao estudo realizado nesse projeto. Os equipamentos de sensoriamento remoto fornecem perfis da atmosfera, enquanto nosso trabalho pretende obter h a partir de dados pontuais, de estações meteorológicas;
- Mesmo uma possível comparação, apesar de equipamentos e metodologias diferentes, não é possível, pois os parâmetros estatísticos são raramente fornecidos, e quando o são, aparecem de forma desconectada e isolada, não sendo realmente representativos.

Os resultados obtidos para o IOA e também r nos mostram que a rede consegue representar de forma bastante promissora o valor de h a partir de dados de superfície, e também outros passíveis de obtenção em estações meteorológicas, o que facilitaria a operação da metodologia. Contudo, cabem ainda maiores investigações, a fim de aprimorar o desempenho da rede. Um fato que parece ser decisivo para a qualidade das estimativas é o número de pares entrada-saída, e principalmente sua limitação de horário. Estão disponíveis apenas dados para 00 e 12 UTC, o que implica em camadas de início da manhã, e camadas já bem desenvolvidas, num geral. A inclusão de dados de outras localidades, os quais pudessem representar o período de desenvolvimento da CLA ao longo do dia, podem trazer informações importantes para o treinamento da RNA, aumentando sua capacidade de generalização e, conseqüentemente, acarretando numa provável melhora dos resultados.

Como dados de superfície utilizamos apenas parâmetros obtidos a partir das medições da radiossondagem, sendo referentes ao primeiro nível de medição reportado, a aproximadamente 1 m da superfície. Contudo, estações meteorológicas de superfície possuem medições de outros parâmetros, como cobertura e tipo de nuvens, radiação e determinação do tempo presente, os quais poderiam também agregar importantes informações ao treinamento e melhorar os resultados.

Por fim, como futuro trabalho, pode-se realizar a operação da rede em outras localidades que não aquelas cujos dados foram utilizados para o treinamento e validação prévias, a fim de avaliar sua capacidade de generalização. Uma proposta seria o emprego da rede já treinada para estimativa de h para dados obtidos pelo LEMMA (Laboratório de Estudos

em Monitoramento e Modelagem Ambiental) da Universidade Federal do Paraná, com valores de h para comparação obtidos a partir de medições provenientes de voos realizados por um Veículo Aéreo Não Tripulado (VANT) [29].

Apêndice A

Parâmetros estatísticos

Tabela A.1: Parâmetros de qualidade para as configurações testadas

Configuração	IOA	IOA _{mod}	r	RMSE	MAD	MBE
st1	0.9406	0.8040	0.8956	437.9921	315.8643	-3.4575
st2	0.9402	0.8007	0.8972	416.8302	300.4384	-12.5887
st3	0.9389	0.8056	0.8904	449.1985	314.8323	-0.5181
st4	0.9432	0.8113	0.9017	426.4814	301.5864	-14.4146
st5	0.9393	0.8041	0.8970	437.2015	310.8407	-15.4081
st6	0.9418	0.8094	0.8975	427.0978	303.1855	-11.1716
st7	0.9417	0.8158	0.8988	429.8743	297.1245	-24.2191
st8	0.9320	0.7928	0.8804	461.9118	329.1498	-6.0460
st9	0.9236	0.7893	0.8710	471.1351	328.2748	-20.8354
st10	0.9373	0.7998	0.8902	430.0635	307.1784	40.6516
st11	0.8873	0.7060	0.8097	559.3019	424.4469	34.3531
st12	0.9117	0.7498	0.8451	515.3455	388.3473	-46.8368
st13	0.9377	0.8059	0.8905	442.2675	310.1550	-20.5549
st14	0.8926	0.7195	0.8217	523.1814	404.0557	-1.3187
st15	0.9382	0.8090	0.8882	448.1169	315.0190	-7.6596
st16	0.9380	0.8031	0.8915	441.3171	313.7042	-41.7061
st1-ln	0.9318	0.7945	0.8868	438.6382	310.4982	-55.0491
st2-ln	0.9187	0.7823	0.8642	484.5067	337.0454	-99.9847
st7-ln	0.9265	0.7960	0.8694	486.8685	330.9670	-59.1836
st15-ln	0.9339	0.7966	0.8806	454.8938	329.2204	-44.4990
st16-ln	0.9287	0.7961	0.8754	457.9225	310.9851	-55.3233
st1-n	0.5033	0.3926	0.2805	212.0871	147.6293	18.0880
st2-n	0.4072	0.3891	0.2886	287.3470	169.3656	-20.3244

st3-n	0.3606	0.3421	0.1342	268.3278	173.4814	-20.1895
st4-n	0.5425	0.4376	0.3903	211.4722	143.7493	-12.0328
st5-n	0.4644	0.3395	0.2407	186.3857	142.4345	8.7008
st6-n	0.4948	0.3293	0.3492	167.6508	134.0175	17.0350
st1-d	0.8404	0.6362	0.7367	589.5853	472.7176	-56.2438
st2-d	0.8309	0.6235	0.7563	583.4907	470.7550	18.7850
st3-d	0.8504	0.6520	0.7628	611.5532	494.8859	82.5978
st4-d	0.8766	0.6824	0.8005	518.3684	412.3393	46.1580
st5-d	0.8238	0.6151	0.7107	588.3079	471.3296	34.1523
st6-d	0.8504	0.6516	0.7714	576.1424	463.3175	-20.2930
sth1	0.5885	0.4390	0.4787	848.3054	678.2616	-23.3577
sth2	0.5776	0.4742	0.3181	1226.6928	931.5061	-239.3113
sth3	0.9305	0.7958	0.8809	466.8184	324.9235	-21.8958
sth4	0.9391	0.7998	0.8940	429.0692	311.4511	20.6295
sth5	0.9292	0.7930	0.8755	447.3634	313.3654	-18.3706
sth6	0.9250	0.7892	0.8702	474.2461	333.8488	11.9600
sth7	0.9204	0.7752	0.8571	481.6895	349.4739	-1.7621
sth8	0.9137	0.7723	0.8469	511.2148	365.8192	-3.2219
sth9	0.9290	0.7889	0.8731	472.4929	337.4623	7.9441
sth10	0.9335	0.7962	0.8822	461.5054	326.1548	11.0901
sth11	0.9280	0.7890	0.8745	471.5321	332.5455	-17.6729
sth12	0.9125	0.7551	0.8437	524.8003	383.4037	-11.2083
sth1-ln	0.5290	0.4725	0.4093	927.8285	657.8684	-346.5461
sth3-ln	0.9296	0.8002	0.8874	484.6952	332.5134	-101.7958
sth4-ln	0.9019	0.7798	0.8412	568.3762	366.5198	-84.3775
sth10-ln	0.9304	0.7950	0.8885	452.1815	322.0264	-87.3667
sth11-ln	0.9209	0.7911	0.8606	511.1396	341.6813	-34.1614
sta1	0.9333	0.8023	0.8814	463.3715	322.6634	35.8751
sta2	0.9473	0.8194	0.9072	407.4014	288.8027	-19.3970
sta3	0.8959	0.7562	0.8264	594.1192	395.9841	17.4335
sta4	0.8975	0.7327	0.8326	539.7450	395.7840	-39.5027
sta5	0.9331	0.7935	0.8853	446.2744	321.3631	-2.4894
sta6	0.8995	0.7389	0.8280	534.3837	394.1374	-11.1993
sta7	0.8689	0.7050	0.7722	621.4328	439.8782	12.0516
sta8	0.9165	0.7512	0.8604	507.3083	378.9456	-52.7677
sta2-ln	0.8812	0.7862	0.8219	604.5890	343.6978	17.8940
sta1-n	0.4021	0.3239	0.1266	233.2331	168.1041	-11.1127

sta2-n	0.4180	0.3331	0.1861	199.3452	142.8677	38.4041
sta3-n	0.4780	0.3869	0.2522	223.0556	146.9516	-2.1526
sta1-d	0.7970	0.6444	0.6684	756.4033	514.6863	-5.9072
sta2-d	0.8327	0.6415	0.7217	624.5283	490.1128	80.0897
sta3-d	0.8034	0.5925	0.6701	656.9168	520.1658	32.4322

Referências Bibliográficas

- [1] BEYRICH, F. Mixing height estimation from sodar data. *Atmospheric Environment*, v. 31, n. 23, p. 3941–3953, 1997.
- [2] HENNEMUTH, B.; LAMMERT, A. Determination of the atmospheric boundary layer height from radiosonde and lidar backscatter. *Boundary-Layer Meteorology*, v. 120, p. 181–200, 2006.
- [3] WANG, X. Y.; WANG, K. C. Estimation of atmospheric mixing layer height from radiosonde data. *Atmospheric Measurement Techniques*, v. 7, p. 1701–1709, 2014.
- [4] TENNEKES, H.; LUMLEY, J. L. *A First Course in Turbulence*. [S.l.]: The MIT Press, 1972.
- [5] ARYA, S. P. *Introduction to micrometeorology*. [S.l.]: Academic Press, 1998.
- [6] BENKLEY, C. W.; SCHULMAN, L. L. Estimating hourly mixing depths from historical meteorological data. *Journal of Applied Meteorology*, v. 18, p. 772–780, 1979.
- [7] GARRAT, J. R. *The atmospheric boundary layer*. [S.l.]: Cambridge University Press, 1992.
- [8] STULL, R. B. *An Introduction to Boundary Layer Meteorology*. [S.l.]: Kluwer Academic Publishers, 1988.
- [9] DIAS, N. L. *Mecânica da Turbulência - Notas de Aula*.
- [10] STULL, R. B. *Practical Meteorology: An Algebra-based Survey of Atmospheric Science*. [S.l.: s.n.], 2014.
- [11] MÜELLER, B.; REINHARDT, J.; STRICKLAND, M. T. *Turbulence in an atmosphere with a non-uniform temperature*. [S.l.]: Springer, 1995.
- [12] ARYA, S. P. *Air Pollution Meteorology and Dispersion*. [S.l.]: Oxford University Press, 1999.

- [13] LI, W. Stable boundary layer height parameterization: Learning from artificial neural networks. *Atmospheric and Climate Sciences*, v. 3, p. 523–531, 2013.
- [14] SEIDEL, D. J.; AO, C. O.; LI, K. Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis. *Journal of Geophysical Research*, v. 115, p. 1–15, 2010.
- [15] ZILITINKEVICH, S. S. Velocity profiles, the dissipation law and the dissipation rate of mean flow kinetic energy in a neutrally and stably stratified planetary boundary layer. 1988.
- [16] ARYA, S. P. S. Parameterizing the height of the stable atmospheric boundary layer. *Journal of Applied Meteorology*, v. 20, p. 1192–1202, 2000.
- [17] PANOFSKY, H. A. *Atmospheric Turbulence - Models and Methods for Engineering Applications*. [S.l.]: John Wiley and Sons, 1983.
- [18] NOAA. *National Oceanic and Atmospheric Administration*. Acessado em 12 de março de 2015. Disponível em: <<http://www.erh.noaa.gov/gyx/UA/introduction.php>>.
- [19] HAYKIN, S. *Neural Networks - A Comprehensive Foundation*. [S.l.]: Prentice Hall International, 1999.
- [20] JORGE, M. P. P. M. et al. Análise preliminar dos dados de sodar relativos à camada limite do pantanal sul matogrossense. *Instituto Nacional de Pesquisas Atmosféricas*.
- [21] SEIBERT, P. et al. Review and intercomparison of operational methods for the determination of the mixing height. *Atmospheric Environment*, v. 34, p. 1001–1027, 2000.
- [22] BEYRICH, F. Mixing height estimation in the convective boundary layer using sodar data. 1994.
- [23] LAMMERT, A.; BÖSENBERG, J. Determination of the convective boundary-layer height with laser remote sensing. *Boundary-Layer Meteorology*, v. 119, p. 159–170, 2006.
- [24] ZELAYA-ÁNGEL, O.; TOMÁS, S. A.; SÁNCHEZ-SINENCIO, F. Atmospheric boundary layer height calculation in méxico city derived by applying the individual eulerian box model. *Bragantia*, v. 59, n. 2, p. 125–137, 2000.
- [25] PELLICIONI, A. et al. Application of neural networks to model the monin-obukhov length and the mixed-layer height from ground-based meteorological data. *Transactions on Ecology and the Environment*, v. 29, p. 1055–1064, 1999.

- [26] FREIRE, L. *Teorias de Camada Limite Atmosférica: modelo de crescimento, fluxo de entranhamento e análise espectral*. Dissertação (Mestrado) — Universidade Federal do Paraná, 2012.
- [27] WILLMOTT, C. J.; ROBESON, S. M.; MATSUURA, K. A refined index of model performance. *International Journal of Climatology*, v. 32, p. 2088–2094, 2012.
- [28] STULL, R. B. *Meteorology for Scientists and Engineers*. [S.l.]: Cengage Learning, 1999.
- [29] MALHEIROS, A. L. et al. Experiências no uso de Veículos Aéreos Não Tripulados para medições de perfis na Camada Limite Atmosférica. *XVIII Simpósio Brasileiro de Recursos Hídricos*.